Computer Networks 182 (2020) 107482

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Real-time dissemination of emergency warning messages in 5G enabled selfish vehicular social networks

Noor Ullah [a], Xiangjie Kong [b,*], Limei Lin [c], Mubarak Alrashoud [d], Amr Tolba [e,f], Feng Xia [g]

[a] School of Software, Dalian University of Technology, Dalian 116620, China
[b] College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China
[c] College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350117, China
[d] Software Engineering Department, College of Computer and Information Sciences, King Saud University, Riyadh 11437, Saudi Arabia
[e] Computer Science Department, Community College, King Saud University, Riyadh 11437, Saudi Arabia
[f] Mathematics and Computer Science Department, Faculty of Science, Menoufia University, Shebin-El-kom 32511, Egypt
[g] School of Engineering, IT and Physical Sciences, Federation University Australia, Ballarat, VIC 3353, Australia

ARTICLE INFO

Keywords:
Emergency Warning Messages
Autonomous Moving Platform
Selfishness
Vehicular Social Networks
Artificial Intelligence

ABSTRACT

This paper addresses the issues of selfishness, limited network resources, and their adverse effects on real-time dissemination of Emergency Warning Messages (EWMs) in modern Autonomous Moving Platforms (AMPs) such as Vehicular Social Networks (VSNs). For this purpose, we propose a social intelligence based identification mechanism to differentiate between a selfish and a cooperative node in the network. Therefore, we devise a crowdsensing based mechanism to calculate a tie-strength value based on several social metrics. Moreover, we design a recursive evolutionary algorithm for each node's reputation calculation and update. Given that, then we estimate each node's state-transition probability to select a super-spreader for rapid dissemination. In order to ensure a seamless and reliable dissemination process, we incorporate 5G network structure instead of conventional short range communication which is used in most vehicular networks at present. Finally, we design a real-time dissemination algorithm for EWMs and evaluate its performance in terms of network parameters such as delivery-ratio, delay, hop-count, and message-overhead for varying values of vehicular density, speed, and selfish nodes' density based on realistic vehicular mobility traces. In addition, we present a comparative analysis of the performance of the proposed scheme with state-of-the-art dissemination schemes in VSNs.

1. Introduction

With the advent of modern vehicular communication technology, the safety of both drivers and passengers has been tremendously improved over the past two decades through the introduction of Intelligent Transportation System (ITS), Vehicular Ad hoc Networks (VANETs), Internet-of-Vehicles (IoV), and many other similar networking frameworks. Various safety-related mechanisms such as collision avoidance systems, dissemination of emergency warning messages (EWMs), mobility traces trajectory-based recommender systems, and other similar proposed schemes have indicated the astounding positive impact of vehicular communication on modern day road traffic [1–3]. However, almost all such proposed schemes operate in multi-hop dissemination mechanisms due to various limitations such as short communication range of the devices and limited network resources under conventional network standards such as Dedicated Short Range Communication (DSRC) and 3G/4G cellular networks. This in turn leads to even more severe issues such as, network congestion due to broadcast storm and hidden terminal collision due to un-reliable wireless connectivity [4]. Therefore, in order to solve the aforementioned issues, Socially Aware Networking (SAN) based novel and innovative solutions have been proposed lately, such as Social Internet of Vehicles (SIoVs) [5] and Vehicular Social Networks (VSN) [6,7] based message forwarding and dissemination schemes. Such schemes integrate human-centric social features with conventional networking solutions using social network theory for message dissemination in the network instead of relying only on connection oriented direct wireless communication [8–11].

Although these schemes have eliminated most of the aforementioned issues but they still face several serious challenges such as algorithm complexity, node detection issues, security risks due to fake alerts, uncertain delivery in erratic and harsh environments,

* Corresponding author.
E-mail address: xjkong@ieee.org (X. Kong).

https://doi.org/10.1016/j.comnet.2020.107482
Received 29 May 2020; Received in revised form 11 July 2020; Accepted 7 August 2020
Available online 14 August 2020
1389-1286/

transmission delay, resource utilization, and lack of standard communication frameworks. Therefore, new solutions and communication frameworks such as 5G enabled information diffusion and dissemination [12,13], Artificial Intelligence (AI) and deep learning based detection networks [14], intelligent grid and Autonomous Moving Platforms (AMPs) based on vehicular clouds and fogs have a great potential [15,16] to solve such issues if integrated with VSNs. However, issues such as the presence of selfish nodes in a network still pose a great threat in successful and timely dissemination of EWMs as selfish behavior can be considered analogous to hidden terminal in conventional wireless network [17]. A selfish node either pushes only its own messages in the network or those which have a potential benefit for it while discarding all other messages and hence refuses to forward or re-route other users' messages in the network [18]. Such nodes can be of two types: (1) purely selfish nodes, which only push their own messages in the network; (2) socially-selfish nodes, which re-route only those messages which have some social benefit for them. Several efforts have been made in Mobile Social Networks (MSNs) to identify selfish nodes and mitigate their effect on the network using various social network and game theory-based methods [19–22]. However, most of these schemes do not consider the highly dynamic nature and network structure of VSNs and AMPs along with the requirements for rapid EWMs' delivery in the network and the impact of delayed dissemination [23]. In addition, the issue of false alarms to avoid any malicious activity in the VSN needs attention and prompt action. Therefore, novel and robust communication platforms such AI based solutions and 5G enabled communication need to be explored for rapid and reliable dissemination of EWMs in a VSN based environment.

Motivated by the necessity to fill the aforementioned gaps, in this paper we propose a 5G enabled VSN based dissemination scheme for EWMs in order to reduce the impact of selfishness as well as encouraging the cooperation among nodes using crowdsensing and social intelligence based incentive scheme. Moreover, the proposed scheme ensures the generation and dissemination of authentic EWMs to avoid security risks associated with false alarms. Therefore, the main contributions of the proposed paper are as follows:

- We devise a predictive mechanism to identify selfish nodes in VSNs by social-intelligence and crowdsensing techniques using their spatio-temporally evolving social ties with the help of 5G enabled vehicular technology.
- In order to achieve real-time dissemination of authentic EWMs with minimum overhead, we design a recursive evolutionary algorithm to identify potential super-spreaders in the VSN by estimating nodes' state-transition probability using a novel propagation model.
- Finally, we perform extensive experimental evaluation to analyze the performance of the proposed scheme under significantly high density of selfish nodes at a varying range of vehicular speed and density in comparison with few well-known state-of-the-art dissemination schemes.

The rest of the paper is organized into the following sections: Section 2 provides literature review of the related work; Section 3 explains the basic network structure and model of the proposed scheme; Section 4 provides a detailed explanation of the proposed scheme; Section 5 explains the operation of the proposed scheme along with the design of the dissemination algorithm; Section 6 gives a detailed and comparative analysis of the experimental results; Section 7 concludes the paper along with identifying future directions.

2. Related works

Recently, several efforts have been made to address the issues associated with dissemination of messages in vehicular networks. For instance, in [9], a social distance metric between nodes has been introduced based on the geographical and social features of nodes in VSNs for data dissemination. In [24], the degree centrality of nodes has been explored to select the most suitable next forwarder in a highway scenario. In [10], a two-pronged strategy based on social interests, groups, and friendships has been proposed to deal with the broadcast storm and hidden node problems in VSNs. Similarly, in [25], a solution for broadcast storm problem has been proposed using a game theoretic model for dissemination. In [7], a detailed analysis of content dissemination schemes in VSNs has been provided. However, these schemes do not address the issues of reliability and authenticity of messages.

To deal with trust and reliability issues, in [8], a social degree metric is calculated based on social activity of a node in the network to estimate the probability for efficient selection of a trusted node as a next forwarder, whereas in [26], a detailed analysis of the proposed scheme has been provided. Similarly, in [11], a trust-based dissemination scheme for EWMs has been proposed using percolation centrality, interaction, and contribution-based utility metric for nodes to evaluate the reputation of nodes and assign them trust score in VSNs. Moreover, a user-posts matrix-based mechanism has been devised to ensure dissemination of authentic EWMs. In [27], a token-based message authentication mechanism has been designed using social cooperation of nodes to reduce the authentication overhead. Similarly, in [28], a context aware reinforcement evaluation system has been proposed to mitigate the influence of malicious nodes and trust establishment in vehicular networks. However, these schemes consider an overall cooperative environment and do not consider the effect of selfish nodes on the dissemination process in VSNs.

In [17,29], severe adverse effects of selfish nodes' presence on data forwarding and network performance have been extensively investigated. Similarly, in [30], a detailed analysis has been provided on non-cooperative behavior of human centric nodes in MSNs and their impact on the network. To solve such issues, several mechanisms have been proposed in mobile and opportunistic networks. For instance, in [22], the reasons for selfish behavior of nodes are identified by eliminating which, data forwarding is enhanced in mobile opportunistic networks. In [21], community acquaintance and intra-community interaction among nodes have been explored in devising an incentive mechanism to encourage cooperation among nodes in the VSNs. In [31], a social contribution-based routing mechanism has been designed to reduce to effect of selfish nodes in vehicular networks. In [18], a social belief and distance-based signaling mechanism has been designed for uncertain data delivery and selfishness in mobile social networks. Similarly, in [19], a game theoretic approach has been adopted to differentiate between individual and socially selfish nodes in MSNs in order to incentivize the data forwarding process. In [32], a bio-inspired incentive mechanism has been proposed using social willingness and depth of social relationships among nodes for ad hoc social networks. Similarly, in [20], a social and non-social credit-based scheme has been devised to stimulate individual and socially selfish nodes for participation in message relaying in socially aware networks.

3. Network model description

3.1. Physical network model

VSNs offer a hybrid network model such that nodes can communicate with mobile stations such as vehicles and pedestrians, as well as with fixed stations such as communication towers and Road Side Units (RSUs). Since previous dissemination schemes have shown that hybrid model enhances the message delivery process, therefore we consider a similar model in this paper. However, in the proposed model, each vehicle is supposed to be equipped with On-Board Units (OBUs) operating under 5G cellular network standards as described in [12]. Similarly, each node is equipped with Global Positioning System (GPS) device to share real-time location, velocity, and time-stamps with the system. Moreover, RSUs are supposed to be installed at given points of the road.

If the nodes are able to communicate range of each other then they are able to share their social and physical information via Vehicle-to-Vehicle (V2V) communication mode, otherwise they can use the facility of Vehicle-to-Infrastructure (V2I) communication mode similar to [13]. Each node's information is shared with RSUs upon contact and is stored for future use, whereas RSUs regularly share the nodes' reputation and social utility information in the network for seamless operation. The memory and storage constraints are not assumed in the proposed model due to enough storage availability on OBUs and data-offloading facility at RSUs as well as to reduce the complexity of the system.

3.2. Social network model

In this paper, we consider VSN as a spatially and temporally evolving social graph such that $G^{(s,t)} = (V, E, (s,t))$, where $V = \{v_1, v_2, \dots, v_N\}$ represents a set of vertices (vehicles or nodes) and $E = \{(i,j) | 1 \le i,j \le N\}$ is a set of edges for all $i,j \in V$, which relate to the social relationships among the nodes. Moreover, we assume that nodes i,j may interact with each other at different physical locations represented by set $PL = \{s_1, s_2, \dots, s_z\}$, and different intervals of time t. All such interactions are captured in a set of events $E_i = \{\epsilon_{i_1}, \epsilon_{i_2}, \dots, \epsilon_{i_n}\}$, for a node i, considered as the observer. Moreover, we organize the social features such as social affiliations, friendships, interaction networks, and interests of each node in a set $SF = \{f_1, f_2, \dots, f_N\}$, such that $f_i = \{f_{i_1}, f_{i_2}, \dots, f_{i_n}\}$ represents features of node i. Another important aspect of a VSN is the presence of various social groups based on the aforementioned social features of the users, therefore similar to [11], we maintain a set of groups $G_s = \{g_1, g_2, \dots, g_n\}$ based on features described in SF. Moreover, as in any online social network, users' feedback plays a vital role in the development of reputation network and incentive mechanism, we consider such a model that a credibility network $C_r = \{c_{r_1}, c_{r_2}, \dots, c_{r_N}\}$ is established based on users' feedback on shared posts and interactions.

3.3. EWM and posts

In this paper we do not assume any constraints on buffer size due to enough storage capacity of the RSUs and OBUs in VSNs. We consider that each node i is capable of creating a set of labeled posts $Post = \{p_1, p_2, \dots, p_n\}$, about a variety of social topics maintained in a temporally evolving weighted label vector space $L_t = \{l_1, l_2, \dots, l_k \mid l_i \in [0,1]\}$ for $1 \le i \le k$. Similarly, to monitor and identify the user-post activity, we maintain a user-post engaging matrix $E_{(t)}^{N \times L} \in \{0,1\}$. Each post comprises of the following attributes: (1) unique Identification Number of the post p_{ID}; (2) the source ID, $p_{src_{ID}}$; (3) the destination's ID, $p_{des_{ID}}$; and (4) the TTL value of the post, p_{TTL}. However, EWM being a special alert beacon contains a single topic of an emergency warning with $l_i = 1$. Other attributes of EWM are: (1) Source Vehicle's ID ID_{src}; (2) Time-stamp of the EWM generation t_{gen}; (3) Location of the identified emergency and source vehicle $Loc_{(em,v_{src})}$; (4) Information about reported emergency, $Info_{em}$; (6) EWM_{TTL}; and (7) The broadcaster's ID, ID_{Br}.

4. Detailed description of the proposed scheme

Below we present a detailed description of the proposed scheme with Fig. 1 illustrating the block diagram of the architecture of CRD whereas Table 1 lists most of the notations used in theory presented.

4.1. Identification of selfish nodes

In order to mitigate the adverse impact of selfishness on the dissemination of EWMs, firstly, we divide nodes into categories based on their social and temporal behavior in the VSN: (1) selfish nodes and (2) Cooperative nodes. The idea is to encourage healthy social cooperation and contribution in the network. However, nodes in realistic scenarios can attenuate their selfishness degree based on their individual and social benefits [19], therefore we consider an altruistic model to define two types of selfish behavior: (1) discrete selfishness; and (2) social selfishness, which are defined as follows:

Table 1
Notations in the proposed scheme and their description.

Notation	Description
$N_{int(i)}$	Number of Interactions of node i
$sim_{(i,j)}$	Similarity value between nodes i and j
$Con_{i(j,k)}$	Contribution of i
$TS^t_{(i,j)}$	Tie-Strength between nodes i and j
$Inf^t_{(i)}$	Influence of i at time t
$B^t_{(i)}$	Evaluated behavior of i at time t
$U^{ps}_{cum_{v_i(t)}}$	Utility value differentiating selfish nodes
$U^{soc}_{cum_{v_{i(j)}}}$	Utility value identifying cooperative node
$C^t_{r_{(i)}}$	Credibility value of i at time t
R^t_i	Reputation value of i at time t
$F^{\epsilon_n}_{i,j(t)}$	Feedback i receives from j at even ϵ_n
$S^t_{act_{(i)}}$	Social activeness of i at time t
$U^t_{s(i)}$	Total social utility of cooperative node at time t
v_{ss}, v_{br}	Super-spreader node and broadcasting node respectively
$p^t_{ss_{(i)}}$	Probability of i to become to super-spread at time t
$\alpha, \sigma, \eta, \rho, \varphi$	Different control and weight variables
ϑ, θ	Threshold values for $U^{ps}_{cum_{v_i(t)}}$, $U^{soc}_{cum_{v_{i(j)}}}$ respectively

We consider the behavior of a node i as discrete selfishness, when it does not relay other users' post but only pushes its own posts in the network, which can lead to grave consequences especially in emergency situations. In this paper we call nodes with discrete selfish behavior as purely-selfish nodes and identify them by utilizing the concept of user-post engaging matrix, their interaction networks, and mobility data in a VSN. Therefore, a node i's interactions in its local community $com_{loc(i)}$ can be calculated which evolve over time as:

$$N_{int(i)}^{local\ t+\Delta t} = \sum_{j=1}^{X} \sum_{k=1}^{L} a_k(i,j)^t + a_k(i,j)^{t-\Delta t} \ , \tag{1}$$

where $a_k(i,j) = 1$ if they interact with each other and 0 otherwise $\forall j \in com_{loc(i)}$. Similarly, as derived in [33], its number of global interactions in the VSN can be calculated as:

$$N_{int(i)}^{global\ t+\Delta t} = \sum_{j=1}^{X} \sum_{k=1}^{L} a(i,j)^t + a(i,j)^{t-\Delta t} \ . \tag{2}$$

Then using Eqs. (1) and (2), we can obtain a normalized number of total effective interactions for each node i at any given time t as:

$$N_{int(i)}^{total} \mid_t = \alpha . N_{int(i)}^{local} \mid_t + (1-\alpha) . N_{int(i)}^{global} \mid_t \ , \tag{3}$$

where α is a control variable. Furthermore, a node's contribution in terms of facilitating other nodes can be calculated as:

$$Con_{i(j,k)} \mid_t = \frac{\sum_{j \neq k}^{N-1} N_{r_{i(j,k)}}}{N_{pi} + \sum_{j=1}^{N-1} N_{s_{j(k)}}} \ , \tag{4}$$

where $N_{r_{i(j,k)}}$ is the number of posts i received from j and $N_{s_{i(j,k)}}$ is the number of posts i sent to j. For instance, if i is connected to nodes j and k but j and k are not directly connected, then in such a scenario, if j sends a message to k via route of i, then i facilitates j. Therefore, this factor is of central importance in contribution and overall utility of the node j. Since vehicular nodes are very dynamic and a node can behave selfishly in a particular environment or group while cooperative in other scenarios therefore, keeping it in mind we calculate a cumulative post-utility score for each node over a given interval of time t as shown

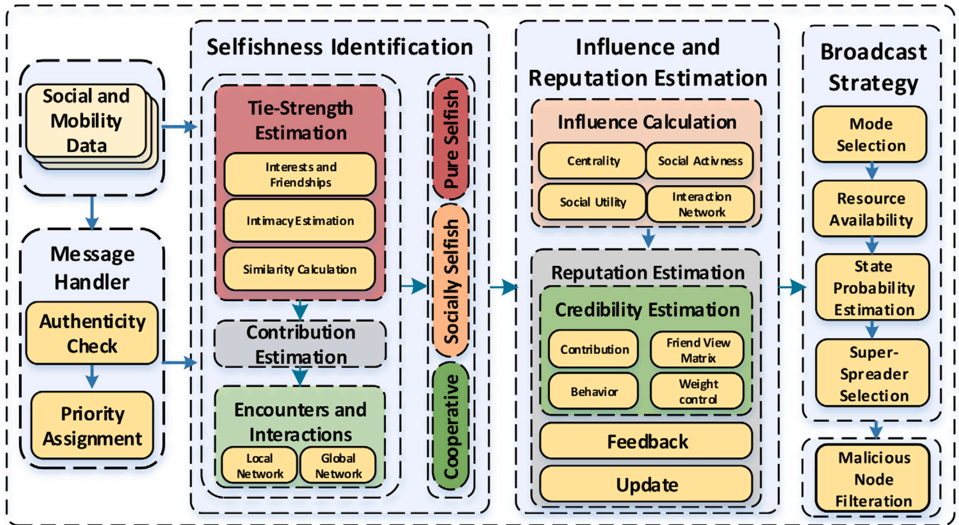


Fig. 1. Block diagram of the architecture of CRD.

in Eq. (5) such that $U^t_{cum_{v_i}} \leq \vartheta$ will determine a node as purely-selfish:

$$U^{ps}_{cum_{v_i(t)}} = \frac{1}{Nt}\sum_{t=0}^{t} e_{ij_{(t)}}.\frac{N^{total}_{int(i)_{(t)}}}{Con_{i(j,k)_{(t)}}} , \tag{5}$$

Similarly, we define social selfishness as the behavior of nodes, when a node pushes its own posts into the network but only relays posts of other nodes based on their social benefits, then such a behavior is termed as social selfishness. The social benefit can be considered in terms of social ties, interests and other social features. Among these features, it has been observed that social tie-strength is of pivotal importance in calculation of a social selfishness calculation [19]. However, identifying the factors that affect the tie-strength (TS) the most has been a hot topic in social network research for a long time [34]. Contrary to [19], where TS is more reliant on similarity between nodes, we found out that only similarity metric may not indicate TS truly with evolving nature of VSNs. Therefore, other factors such as community activity, information about interaction networks, duration of interactions, mobility data, and aging factors must be considered in calculating a TS value for each node. Generally, TS can be calculated as:

$$TS = \beta_0 + \sum_{i=1}^{n} \beta_i X_i , \tag{6}$$

where β_o is a seed value and X_i represent a specific social factor. In this paper we consider three main factors which are: (1) x_1, similarity between nodes i and j; (2) x_2, Intimacy between nodes by considering the nodes' activity and exchange of posts among their connection over a given amount of time; (3) x_3, community and common mobility information. Therefore, calculating each factor requires information over an extended period of time in a VSN environment between any two nodes for accurate identification. For this purpose, we consider the first factor as $x^t_1 = sim^{t+\Delta t}_{(i,j)}$ in Eq. (7), which is calculated based on *homophily-theory* as derived in [11], therefore:

$$x^t_1 = \begin{cases} sim^{f_i \cup f_j}_{(i,j)}, & \text{at } t = 1 \\ sim^{t-\Delta t}_{(i,j)} + \delta^2 . \sum_{i=1}^{t-\Delta t} sim^t_{(i,j)}, & \text{otherwise,} \end{cases} \tag{7}$$

where δ is a constant to select the priority of a time-stamp by varying its exponent, whereas $sim_{(i,j)}$ represents similarity between nodes i and j. In order to calculate x_2, we take into account the interactions of node i with other nodes in its community along with the rest of the nodes in the network and the exchange of posts among them. As a rule of thumb in social theory the duration of each interaction indicates the level of intimacy between two nodes. Therefore, we consider the contact duration along with aging factor of such interactions because the recent interaction data provides better indication of a node's selfish behavior. Therefore, we calculate x_2 as:

$$x^t_2 = \Big(\frac{m.N^{total}_{int(i)_{(t)}}}{\sum_{int=1}^{m} \log D_{int(i,j)}} + \frac{\sum_t \sum_{i,j\in V}^{N} sim(l_p \mid^y_{i=1}, l_p \mid^z_{j=1})}{l_{p(i)} + l_{p(i)}}\Big).e^{-3t} , \tag{8}$$

where $D_{int(i,j)}$ indicates the duration of interaction between nodes i and j. To calculate the third factor x^t_3, in this paper, we consider the number of common groups between nodes i and j along with their common connections or friends and the common locations that both the nodes have visited over a given period of time. We realize this by these quantities into a combined factor x^t_3 as:

$$x^t_3 = \sigma.g(i,j) \mid_t + \eta. \sum_{t=0}^{t} f_{r_{(i)}} \cap f_{r_{(j)}} + \rho.N_{comLoc(i,j)} , \tag{9}$$

where σ, η, and ρ are user controlled variables with values ranging between 0 and 1 in order to assign weight to each factor in above equation and $N_{comLoc(i,j)}$ represent the number of common visited locations. For the purpose of simplicity, we consider $\sigma = \eta = \rho = 0.3$ in our experiments. Then by having all the factors in Eqs. (7), (8), and (9), we can calculate the real-time log-normal TS value between nodes i and j for each time instance as:

$$TS^{t+\Delta t}_{(i,j)} = \begin{cases} \frac{1}{10}\log(\beta_0 + \sum_{i=1}^{3} \beta_i.X_i), & \text{at } t = 1 \\ TS^t_{(i,j)} + \delta^2.(TS^{t-\Delta t}_{(i,j)}), & \text{otherwise.} \end{cases} \tag{10}$$

Then using Eqs. (5) and (10), we can derive a cumulative social utility value for each node i at each time interval t to determine if it should be considered as socially-selfish or cooperative node in the VSN as follows:

$$U^{soc}_{cum_{v_{i(j)}}} \mid_t = \alpha.TS^{t+\Delta t}_{(i,j)} + (1-\alpha)U^{t+\Delta t}_{cum_{v_i}} , \tag{11}$$

where $0 \leq \alpha \leq 1$ is a user defined control variable to assign weights to each factor in Eq. (11). If $U^{soc}_{cum_{v_{i(j)}}} > \theta$, where θ is a system assigned threshold value, then the node is considered as cooperative node and selfish otherwise.

4.2. Influence and reputation calculation

Since higher number of broadcast messages negatively impacts the network performance, therefore, the proposed scheme focuses on

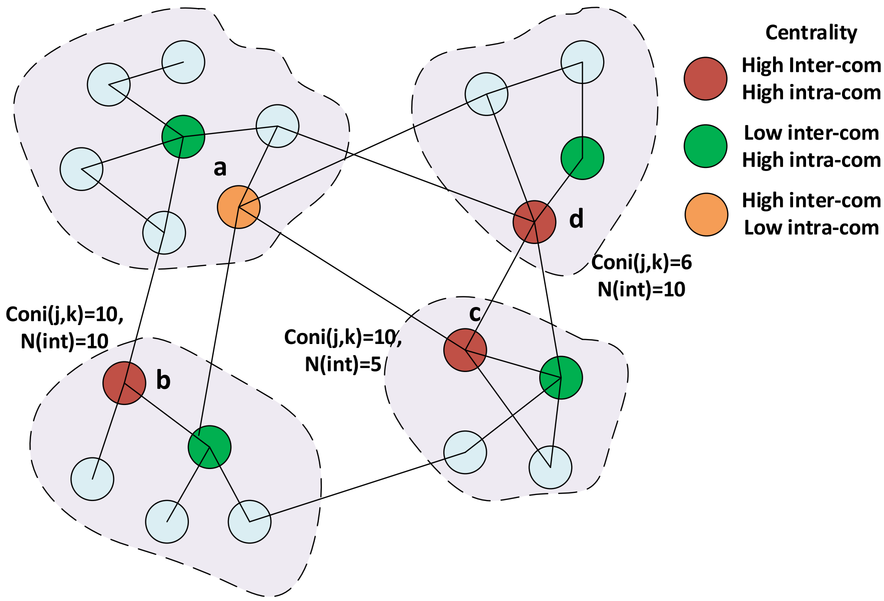


Fig. 2. Example of influential node's selection based on social activity.

identifying key influential nodes to be considered as super-spreaders in the network. For this purpose, we utilize the percolation centrality metric along with $U^{soc}_{cum_{v_{i(j)}}}$ to evaluate the influence of a node i at any given time t as:

$$Inf^{t+\Delta t}_{(i)} = \begin{cases} \frac{\alpha.U^{soc^t}_{cum_{v_{i(j)}}} + (1-\alpha).Per^t_{c(i)}}{\max_{j\in V} Per^t_{c(j)}}, & \text{at } t = 1 \\ Inf^t_{(i)} + \delta^2.Inf^{t-\Delta t}_{(i)}, & \text{otherwise,} \end{cases} \tag{12}$$

where $Per^t_{c(i)}$ represents the percolation centrality of a node i at a given time t. As discussed earlier that nodes in VSNs have very high mobility, therefore, a VSN can be treated as dynamic network where nodes' properties may rapidly change both spatially and temporally. Moreover, EWM's dissemination is a spreading phenomenon having a tendency to percolate from one state to another. To explain this, let us consider a scenario where a source node $i \in V$, undergoes a change in state from s^t_i to s^t_j upon initiation, representing the condition that whether the node is informed or not? For instance, if we divide nodes into states s_1 and s_2 where s_1 means that a specific node is uninformed yet and s_2 represents that it has been informed. Then, such a change of state over time may also impart a change of state in the target vehicles from states s^t_i depending on their previous states. Therefore, we consider percolation centrality, $Per_{c(i)}$ to identify influential nodes in VSN for dissemination of EWMs at any time t [11]. Fig. 2 illustrates the selection process of high utility node. For instance, node a is highly connected among the four communities but is least connected in its own community. On the other hand, b is highly connected locally but has less global connectivity with high contribution and interactions. Similarly, nodes c and d have both high local and global connectivity but with different contribution and interaction levels. Among these nodes, d is has the higher likelihood to be the most influential node in the network based on its high connectivity, interactions, and contribution in the network. In order to calculate the reputation of the nodes in the VSN we follow the model of [11] with slight modifications such that a new temporally evolving behavior metric $B_{(i)_{(t)}}$ is calculated for each node i in the VSN based on its $U^{ps}_{cum_{v_{i(t)}}}$ value combined with its shared posts and their reviews over an extended period of time t as:

$$B^t_{(i)} = \mu.U^{ps}_{cum_{v_{i(t)}}} + (1-\mu)\left(\frac{\sum_{j\in V}\sum_{k-1}^{n} \mid R_{pos_{k(j)}} - R_{neg_{k(j)}} \mid}{\mid P_i \mid}\right), \tag{13}$$

where $R_{pos_{k(j)}}$ and $R_{neg_{k(j)}}$ represent positive and negative reviews for the k_{th} post of node i by node j respectively, and μ is the weight control factor which can be calculated as:

$$\mu = \frac{\sum_{j\neq k} g_{j,k}(i) \mid_t}{N^{total}_{int(i,j)} \mid_t + N^{total}_{int(i,k)} \mid_t}, \tag{14}$$

where $g_{j,k}(i)$ represents number of geodesics between nodes j and k that via node i. Using Eqs. (4) and (13), we can obtain the credibility metric for each node at time t as:

$$C^{t+\Delta t}_{r_{(i)}} = \begin{cases} \mu.Con^t_{i(j,k)} + (1-\mu)B^t_{(i)}, & \text{at } t = 1 \\ C^t_{r_{(i)}}\delta^2 + C^{t-\Delta t}_{r_{(i)}}, & \text{otherwise.} \end{cases} \tag{15}$$

Consequently, a temporally evolving dynamic reputation value can be obtained based on aforementioned calculated credibility value of a node and the general feedback it receives on its posts from the rest of the nodes as:

$$R^t_i = C^t_{r(i)} \times \sum_{j\neq l}^{N-1} F^t_{(i_{(j)})}, \tag{16}$$

where $F^t_{(i_{(j)})}$ is the feedback received by i in each interaction which is recursively called in each iteration of R^t_i and is calculated based on reviews of its friends as well as from the rest of the members of VSN as:

$$F^{\epsilon_i}_{i,j_{(t)}} = U^t_{s(i)} \times E^t_{fr(i)} + (1 - U^t_{s(i)}) \times R^t_i, \tag{17}$$

and $E^t_{fr(i)} = (A^T)^K.az_i$ is the evaluation of node i based on the feedback of its friends at each event ϵ_i, in order to mitigate the potential negative reviews by malicious nodes whereas $U^t_{s(i)}$ provides the temporally evolving weightage to each review and is calculated in Section 5 using Eq. (21). Therefore, we take the feedback of node i's friends about another node j based on which it evaluates node j and is stored in a matrix $Z^{X\times N-1}$ such that $z_{ij\in[0,1]}$. Therefore, we can derive a matrix $A^{X\times K}$ containing all the feedback values for each post of nodes similar to [35] as:

$$A = \begin{bmatrix} a^1_1 \cdot z_{11} & a^1_2 \cdot z_{12} & \dots & a^X_1 \cdot z_{1K} \\ a^2_1 \cdot z_{21} & a^2_2 \cdot z_{22} & \dots & a^X_2 \cdot z_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ a^X_1 \cdot z_{X1} & a^X_2 \cdot z_{X2} & \dots & a^X_X \cdot z_{XK} \end{bmatrix}, \tag{18}$$

where $a^j_i = 1$ if i declares j as credible and 0 otherwise. Moreover, we normalize the matrix as $az_{ij} = \frac{a^j_i.z_{ij}}{\sum_n a^j_i.z_{jn}}$ if $a^j_i.z_{ij} \neq 0$ and 0 otherwise.

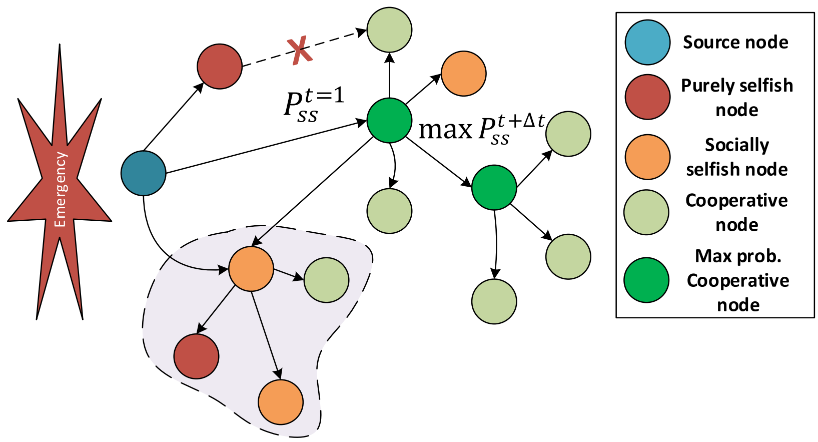


Fig. 3. Super-spreader selection and dissemination process of CRD.

Due to dynamic nature of VSN, we update the $R^t_{(i)}$ value with temporal evolution at a decaying rate τ as:

$$R^{t+\Delta t}_{i,j} = e^{-(\tau)} R^t_{i,j} + (1-\tau) R_{imp_i} \;, \tag{19}$$

where R_{imp_i} is an impact function as derived in [36].

5. Operation of dissemination algorithm

In order to efficiently disseminate the EWM in the VSN with minimum message overhead, transmission delay, and high authenticity, this paper proposes a multi-step strategy-based scheme Cooperative Real-time Dissemination (CRD), for which the pseudocode is given in Algorithm 1. In the first step, the EWM authenticity is checked using the method proposed in our previous work [11] by incorporating a heterogeneous user-post interaction network and their engaging matrix E as described in Section 3.3. For this purpose, we treat a false alarm as a fake news in a social network, and hence adopts a fake news detection mechanism. Therefore, we use the data collected from user-post interaction network to identify the credibility of nodes as detailed in [11]. As a general trend, low credibility nodes are more likely to raise a false alarm than the higher ones. Hence, by using the Non-negative Matrix Factorization (NMF) method, the document-word matrix can be exploited to extract a non-negative matrix that resent the shared posts. Similarly, the adjacency matrix $A^{N\times N} \in [0,1]$ is then used to obtain another non-negative matrix that represents the users of the shared posts. Moreover, shared partially labeled posts stored in L_t along with the vehicle-post engaging matrix are then used to derive equations as given in [11], which are used to infer the authenticity of EWM. Once the authenticity of EWM is acknowledged, it is accredited as EWM with a unique accredited broadcast ID to make it priority message in the posts' header and is broadcast in the communication range of source vehicle.

In the second step, the proposed algorithm collects data of the nodes in the concerned area of the given VSN to identify the behavior of nodes as purely-selfish, socially-selfish, or cooperative. Since, purely-selfish nodes disrupt the broadcasting process, the proposed scheme discourages such nodes in dissemination process and are blocked to be considered as broadcaster immediately but are treated as conventional nodes to be informed only. However, socially-selfish nodes can still be of importance as they can spread the EWM among its close peers and communities to improve its social utility. Therefore, if a node is identified as socially-selfish then the most influential node is selected among the clusters of such nodes for broadcasting the EWM among its peers using Eq. (12).

In the third step, the cooperative nodes are considered for potential super-spreader role. Therefore, to identify such a node, many factors must be taken into account such as: how much is the reach of the node in terms of its direct and indirect connections?; how much is it active in communicating via posts and interactions within the communities and in the VSN generally?; and how closely is the node connected with specific nodes rather than a general population of the VSN?. To realize these factors, similar to [33] but with slight improvements, CRD first evaluates the degree of social activeness of each node i in the VSN within a given amount of time as:

$$S^t_{act(i)} = \frac{N^{total}_{int(i)}|_t}{N-1} \times \left(\frac{N^{t-\Delta t}_{nbr_{(i)}} \cap N^t_{nbr_{(i)}}}{N^{t-\Delta t}_{nbr_{(i)}} \cup N^t_{nbr_{(i)}}}\right) , \tag{20}$$

where $N_{nbr_{(i)}}$ represent the number of neighbors of node i having the least social-distance in the given scenario. Similarly, the tie-strength for node i is obtained using Eq. (10), such that a higher tie-strength with certain specific nodes impacts the performance negatively as priority are given to the highly intimate nodes. Therefore, in case a node is being identified as cooperative node, then the Proposed Scheme calculates a total social utility value for each node in the VSN to identify the most influential node v^{act}_{inf} based on the aforementioned equations as:

$$U^t_{s(i)} = \varphi . Inf^t_{(i)} - \frac{(1-\varphi S^t_{act(i)}).\sum_{j\in V} TS^t_{(i,j)}}{S^t_{act(i)}} \;. \tag{21}$$

A node with highest $U^t_{s(i)}$ in the close vicinity of v_{src} is selected as v^{act}_{inf} and first super-spreader of EWM in the first step of broadcast. Therefore, it broadcast the EWM to all nodes in its communications and all its peers in local communities and VSN. Since VSNs exhibit small world phenomenon [37,38], therefore to avoid the flooding of EWM in the network, the proposed scheme identifies the highest reputed nodes that have the least value of geodesics from v^{act}_{inf}. All the distances between v^{act}_{inf} and v_i at any given time t are maintained in a matrix $Dv^{act\;1\times N}_{inf_t}$ whereas, a distance matrix for all nodes in the VSN, i.e., v_i is also maintained as $Dv^{N\times N}_t$. Then we can obtain the shortest paths between v^{act}_{inf} and the most reputed and influential nodes in its influence circle in a given VSN using the distance matrix $D^{N\times N}\;|_t$ such that $d^t_{ij} = \min_{0<n\leq N} dv^{act^t}_{inf} + dv^t_{ij}$. Therefore, nodes $v^t_{br} \in SS$ that have the least social-distance to v^{act}_{inf} with highest $R^t_{(i)}$ are selected and prioritized to be considered as new broadcasters of EWM as:

$$v^{t+\Delta t}_{ss} \leftarrow \{SS \mid (\max_{i\in v^{t-\Delta t}_{br}} R^t_{(i)}, \min d^{t-\Delta t}_{ij})\} \;, \tag{22}$$

where $v^{act}_{inf} = v^{t=1}_{br}$. After successful execution of the broadcast the process is repeated again from second step onwards until all the nodes in the VSN are informed and identified via Ack messages. Throughout the dissemination process each node transitions between one of the four distinct states that are: (1) uniformed s_1; (2) informed s_2; (3) super-spreader s_3; and (4) acknowledged s_4, caused by its activities of posts'

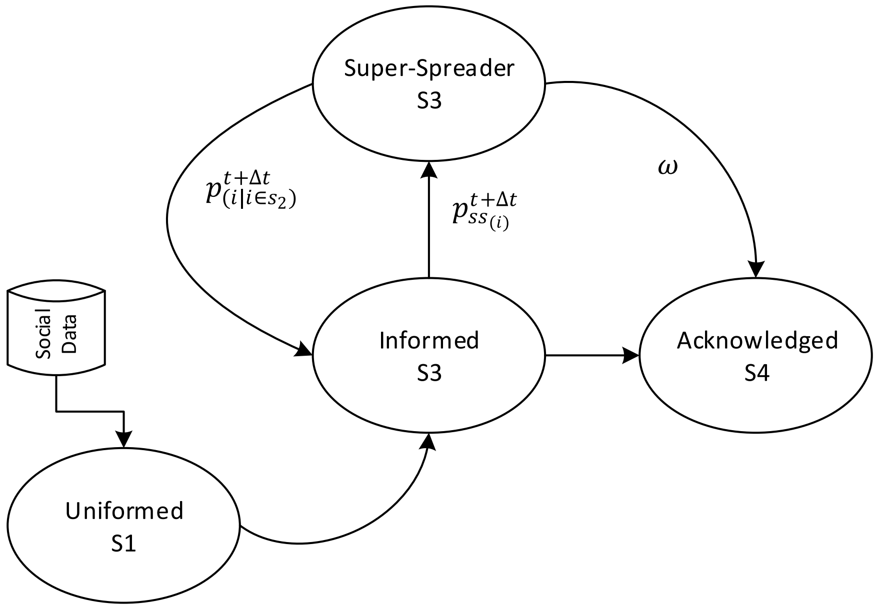


Fig. 4. State transition diagram for CRD.

reception and transmission similar to the meme propagation model described in [39] and its enhanced variants. Initially each node except for v_{src}, is in s_1. Upon reception of EWM, the node either transits into s_2 or s_4. In case the node transits into s_2, then the probability for it to become a super-spreader i.e. transit into s_3 increases based on its $U^t_{s(i)}$, $R^t_{i,j}$, and $U^{soc}_{cum_{v_{i(j)}}}$ values. Therefore, each informed node upon reception of EWM may become a super-spreader with a probability p_{ss} to broadcast EWM in the VSN, or remain in s_2. The probability of a node to be retained in s_2 is thus determined by its previous state, its influence in the network, and $U^{soc}_{cum_{v_i}}$ along with its spatial and temporal positioning in the VSN as:

$$p^{t+\Delta t}_{i|i\in s_2} = 1 - (1 - p_{(i|i\in(s_1\cup s_2))}).(1 - Inf^t_{(i)}.U^{soc}_{cum})^{y-1} , \tag{23}$$

where y is the post counter iteration for each event. Any node in s_2 will eventually transit into s_4 where it only acknowledges the EWM but does not disseminate it further, ensuring that each node disseminate the EWM only once in order to avoid re-broadcasts. Similarly, a node's state transition from s_2 to s_3 i.e. becoming super-spreader can then be estimated as:

$$p^{t+\Delta t}_{ss_{(i)}} = \omega.p^t_{(i|i\in s_2^{t-\Delta t})} + (1-\omega).Inf^t_{(i)} , \tag{24}$$

where ω is the weight factor depending on the nodes' reputation and social utility at a given time t and is calculated as:

$$\omega = \mu_\tau.\frac{U^{soc}_{cum_{v_{i(j)}}}}{R^t_{(i)}} . \tag{25}$$

Therefore, the probability of a node to become a super-spreader $p^{t+\Delta t}_{ss_{(i)}}$ at next time interval gets higher only if it is a cooperative node and has higher reputation in the VSN whereas socially-selfish nodes' probability lowers at each iteration. Any node that has $p^{t+\Delta t}_{ss_{(i)}} < p_{threshold}$ transits to acknowledged state and does not take part in the dissemination in next iteration. Fig. 3 illustrates the dissemination process based on super-spreader selection, whereas, the overall state transition process is shown in Fig. 4. As an example, as shown in Fig. 3, the source node generates a EWM and broadcasts it in its communication range as well as to its direct connections. However, among the receiving nodes, only a cooperative node that has highest P_{ss} value is selected as the next hop broadcaster whereas a purely selfish node is assigned s_4 which can only receive and acknowledge the EWM but is not considered for the broadcaster role. Moreover, as socially selfish nodes have strong ties within their closed circles, therefore a socially selfish node is also allowed to attain s_3 in its own community and never preferred over a cooperative node that has same P_{ss} value. This process is carried out until the EWM is propagated in the network based on a set TTL value.

In order to disseminate the EWM efficiently, we consider both centralized and distributed modes of communication depending on the availability of resources. For this purpose, if RSU is in range then V2I mode is enabled where a super-spreader node is selected by the RSU using Eq. (22) among all the receiving nodes which then broadcasts the EWM in its range as well as to the nodes with least distance and higher reputation values and the process is repeated until the TTL limit is reached. On the other hand, if RSU is not in range, then V2V mode is enabled where the receiving node broadcasts the EWM in its range and based on the super-spreader probability using Eq. (24), a node with highest P^t_{ss} is selected as the next broadcaster which further broadcasts the EWM until TTL limit is reached or if RSU is available in the range then V2I mode is enabled again.

6. Experimental results and analysis

6.1. Dataset and simulation setup

In order to generate a realistic scenario and behavior observation, we based our experiments on the realistic crowdsourced dataset provided by Shanghai QiangSheng Holding Co., Ltd. The published data focuses on the mobility and surrounding activity of 13,750 taxis in Shanghai for a month containing detailed trajectory traces about various attributes such as location, speed, angle, warnings, trips to common places, status, and conditions with a unique identity number assigned to each vehicle. However, in this paper our main focus is on high speed vehicles and scarce road conditions similar to a highway, therefore, we extracted mobility data of 1000 vehicles over road sections of 4 kilometers divided in 4 lanes each side that fulfill most of the aforementioned criteria. After performing several operations based on the recommendations in [40,41] for real-world road-network topology generation, we created a mobility dataset for randomly seeded social features such as interests, posts, groups, interactions, and friendships which were evolved over an extended period of time in a controlled environment using the ONE simulator. Moreover, we performed the aforementioned operation on a range of varying vehicular densities, speeds, and selfish nodes' densities to quantify the impact of selfishness delivery ratio, transmission delay, and message overhead on the

Algorithm 1 Pseudocode for the proposed scheme

INITIALIZATION:
if $V_{src} \in V$ detects emergency **then**
  Generate EWM
  if $EWM = P_{False}$ **then**
    validate → Broadcast with Alert Note
    if $\exists V_i \in V \mid \{v_i \rightarrow (\max R_i^t, \min distance_{v_{src}})\}$
    Validates $EWM = P_{False}$
    Terminate, Blacklist V_{src}
  else
    if $t < ttl_{EWM}$ **then**
      if $V_{received} \rightarrow U^t_{cum_{v_i}} < \vartheta$ **then**
        $V_{received} \rightarrow S_4$
      end if
      if $V_{received} \rightarrow U_{cum_{v_i}} and U^{soc^t}_{cum_{v_{i(j)}}} < \theta$ **then**
        Broadcast to $v_{received}$ connections
        $v_{received} \rightarrow s_2$
      else
        if RSU is in Range **then**
          Initiate → V2I Mode
          $V_{received} \in V$, Select → max $v_{ss}^{t+\Delta t}$
          Broadcast → $v_i \ni v_{received}$'s Range
          Broadcast to all ↓
          $v_i \in V \mid (\max_{i \in v_{br}^{t-\Delta t}} R^t_{(i)}, \min d_{ij}^{t-\Delta t})$
          Repeat V2I Mode
        else
          Initiate → V2V Mode
          Broadcast to $v_j \in V$ in range
          Identify $v_k \mid v_k \rightarrow \max p^{t+\Delta t}_{ss_{(i)}}$
          $v_j = v_{ss}^{t+\Delta t}$, Broadcast and Scan RSU
          if RSU is in Range **then**
            Repeat V2I Mode
          else
            Repeat V2V Mode
          end if
        end if
      end if
    end if
  end if
end if

Table 2
Simulation parameters.

Parameter	Value
Vehicular speed	80–120 km/h
Vehicular density	24–120 vehicles/km
Selfish nodes density	10–50%
Node's transmission range	80 m
RSU's transmission range	300 m
CBR	0.2 s
Road length	4000 m
MAC protocol for V2I	IEEE802.11p

performance the proposed scheme. In order to represent a real-world network scenario, we considered selfish nodes' density in the VSN between 10% and 50% because practically no social network can be free of selfish nodes whereas selfish nodes' density above 50% will be an impractical scenario for a network to remain stable [19,30]. The rest of the simulation parameters are listed in Table 2.

6.2. Performance evaluation

To understand the performance and impact of the proposed scheme for the defined social metrics such as Tie-Strength TS and social utility $U^t_{s(i)}$, we evolved the system over an extended period of time and observed its behavior. For instance, Fig. 5 illustrates the evolution of cumulative social utility values $U^t_{cum_{v_i}}$ over time for different values of control parameter δ, which indicates that a higher value of δ results in higher rate of utility gain. Therefore, we selected $\delta = 0.7$ in experiments for general trend towards encouraging the cooperative nodes, which eventually results in higher probability of such nodes' selection as super-spreader. Similarly, Fig. 6 provides the rate of change in probability of cooperative node selection as super-spreader for different values of φ with evolving time, which in turn indicates that a node's social activity and stronger ties play higher role in dissemination process than only its centrality. Moreover, in Fig. 7, we compared the performance of CRD, [11], and [19] for the probability to select a cooperative node instead of a socially-selfish node using the TS values of each scheme. The results showed that the proposed scheme performs better than its competitors with a probability of about (90%–99%).

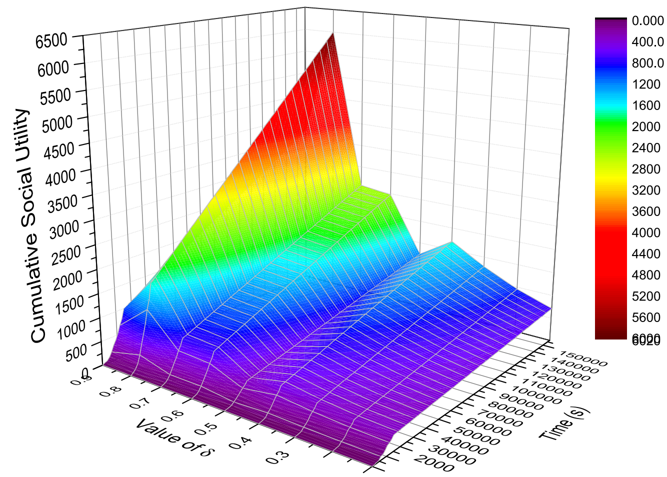


Fig. 5. Effect of different values of δ on social utility.

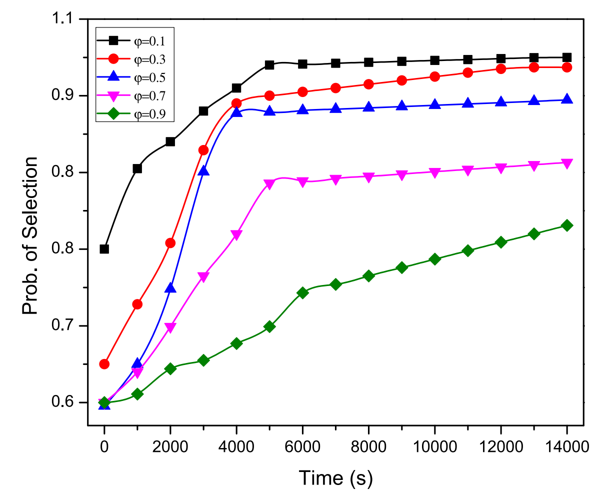


Fig. 6. Effect of φ on probability of suitable node selection.

For further evaluation, we analyzed the experimental results of the proposed scheme in comparison with three well reputed state-of-the-art VSN based dissemination schemes namely SCARF [8], CDF [21], TDS [11] under the aforementioned network simulation environment and parameters for delivery ratio, hop-count, transmission delay, and total number of transmitted messages by varying the densities of vehicles and selfish nodes, and vehicular speed. The reason for selecting the aforementioned schemes is the similar network structure as these are

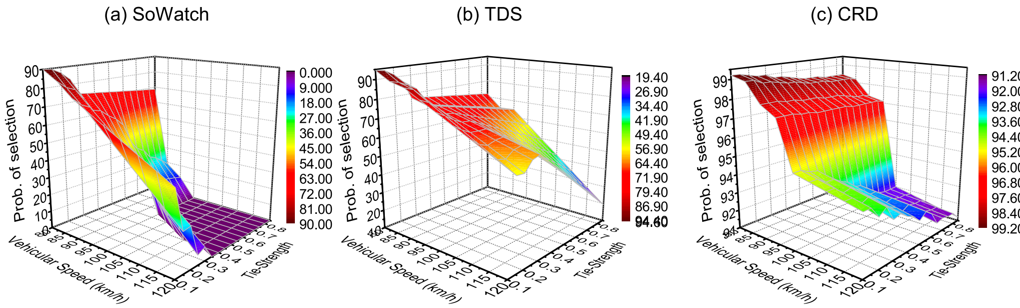


Fig. 7. Comparison of the effect of TS on probability of selfish node's detection.

designed for VSN environment. Moreover, most of these schemes deal with one or more similar issues such as rapid message dissemination, selfishness, ensuring reliability and authenticity, as well employing social properties based mechanism to achieve these goals.

6.2.1. Delivery ratio

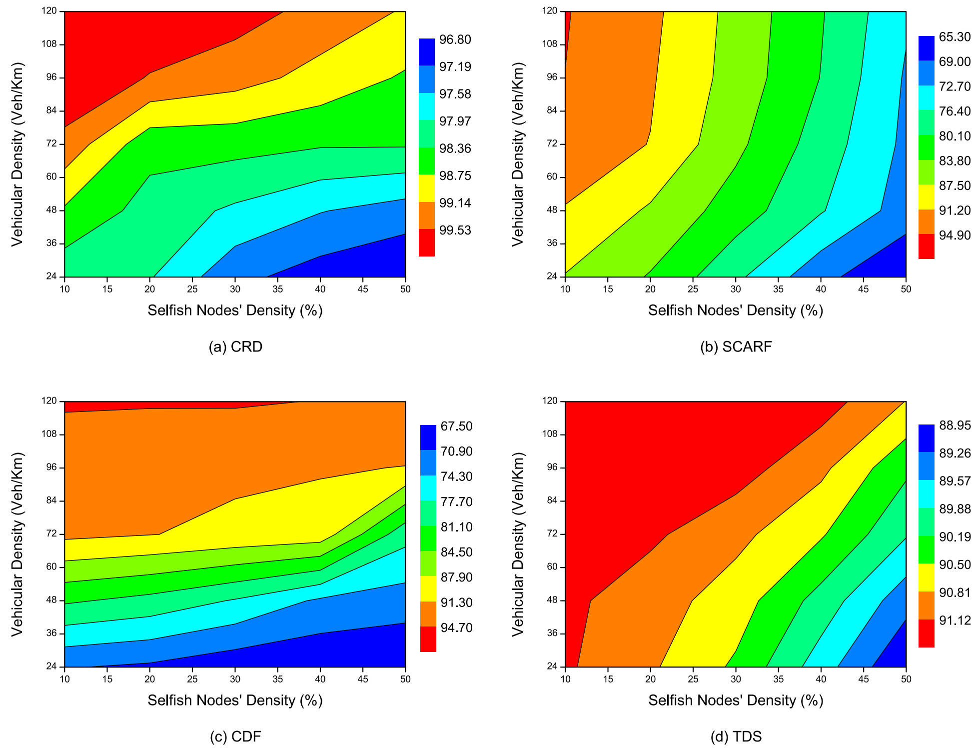
Delivery ratio refers to ratio of number of received vehicles to the total number of vehicles in the VSN. Fig. 8 illustrates the values of delivery ratio over a range of vehicular and selfish nodes' densities in the network at a constant speed. The results indicate that the proposed scheme achieves a delivery ratio of 96.8%–99% for a varying vehicular density of 24–120 vehicles/km which is about 31.5%–4.1% higher than SCARF with 65.3%–94.9%, 29.3%–4.3% higher than CDF with 67.5%–94.7%, and 8.7%–7.9% higher than TDS with 88.1% to 91.1% respectively. These results show that CRD achieves significantly higher delivery ratio than its competitors regardless of the vehicular density under varying selfish nodes' density. Moreover, to evaluate the degree of variation in delivery ratio for each value of vehicular density while varying selfish nodes' density (10% to 50%), Fig. 9 shows that the proposed stays relatively more stable under the aforementioned than the rest with median (97.4–99.6) with a difference of 2.2, variance (0.02–0.2) with a difference of 0.18, co-efficient of variance (0.001–0.004) with a difference of 0.003, and mean square error (MSE) (0.05–0.15) with a difference of 0.1 as compared to SCARF's (78.3–85) with a difference of 6.7, (40.2–56.8) with a difference of 16.6, (0.07–0.09) with a difference of 0.02, and (2.11–2.51) with a difference of 0.4, CDF's (71.3–94.9) with a difference of 23.6, (0.05–19) with a difference of 18.95, (0.002–0.05) with a difference of 0.0048, and (0.07–1.45) with a difference of 1.38, and TDS's (90.4–91.3) with a difference of 0.9, (0.04–0.6) with a difference of 0.56, (0.002–0.008) with a difference of 0.006, and (0.07–0.25) with a difference of 0.18 respectively. These results indicate that CRD not only achieves higher delivery ratio in each scenario but also stays significantly stable than its competitors for a varying range of vehicular and selfish nodes' density.

Fig. 10 illustrates the values of delivery ratio for each scheme under varying vehicular speeds (80%–120%) for a range of selfish nodes' density (10%–50%). The results validate that the proposed scheme achieves the highest delivery (99%) which is approximately 35% (at 80 km/h) and 13% (at 120 km/h) higher as compared to SCARF's (64%–87%), 38.7% and 13.5% than CDFs (60.3%–85.5%), and 24%–8% than TDS's (75%–91%) respectively. These results show that CRD achieves significantly higher delivery ratio than its competitors regardless of the change in vehicular speed under varying selfish nodes' density. In addition, to evaluate degree of variation in delivery ratio by varying the selfish nodes' density from 10% to 50%, Fig. 11 provides the data of the median (99.2–99.7) with a difference of 0.2, variance (0.3–0.6) with a difference of 0.3, co-efficient of variation (0.018–0.002) with a difference of 0.002 and MSE (0.6–0.8) with a difference of 0.2 in comparison to SCARF's (76–77.2) with a difference of 1.2, (65.3–68.7) with a difference of 3.4, (0.105–0.109) with a difference of 0.004, and (2.69–2.76) with a difference of 0.07, CDF's (65.6–88.7) with a difference of 23.1, (0.12–28.7) with a difference of 28.18, (0.004–0.07) with a difference of 0.066, and (0.11–1.76) with a difference of 1.65, and TDS's (85.3–89.1) with a difference of 3.8, (6.26–24.8) with a difference of 18.54, (0.02–0.05) with a difference of 0.03, and (0.8–1.6) with a difference of 0.8 respectively, for a range of vehicular speeds (80 km/h–120 km/h). These results indicate that the performance of CRD remains very stable as compared to its competitors at a wide range of varying vehicular speed and selfish nodes' density in the network.

6.2.2. Transmission delay

Fig. 12 provides results for total time taken by the EWM to spread in the network under varying selfish nodes' density (10%–50%) for varying vehicular densities (24 veh/km–120 veh/km). The results indicate that the proposed scheme outperformed its competitors in terms of transmission delay with values of 3.79–6.1 ms which is significantly less than SCARF's (6.3–16 ms), CDF's (23–52.6), and TDS's (7.6–14 ms) by approximately 39.84%–61.88%, 83.52%–88.4%, and 50.13%–56.43% respectively. These results show that CRD disseminates the EWM much faster than its competitors regardless of the vehicular density under varying selfish nodes' density. Similarly, to evaluate the degree of variation in transmission delay for each value of vehicular density under varying selfish nodes' density from 10% to 50%, Fig. 13 provides the minimum and maximum values of median (3.8–6.2) with a difference of 2.4, variance (0–0.018) with a difference of 0.018, co-efficient of variance (0–0.02) with a difference of 0.02, and MSE (0–0.04) with a difference of 0.04 for the proposed scheme as compared to SCARF's (9.3–12.3) with a difference of 3, (1.23–15.65) with a difference of 14.42, (0.1–0.3) with a difference of 0.2, and (0.3–1.3) with a difference of 1, CDF's (24.7–37) with a difference of 12.3, (3.53–43.95) with a difference of 40.42, (0.07–0.16) with a difference of 0.09, and (0.6–2.2) with a difference of 1.6, and TDS's (8.6–13.5) with a difference of 4.9, (0.17–0.87) with a difference of 0.7, (0.04–0.08) with a difference of 0.04, and (0.13–0.32) with a difference of 0.19 respectively. These results indicate that the proposed scheme is least affected by variation in vehicular and selfish nodes' density in terms of transmission delay as compared to its competitors.

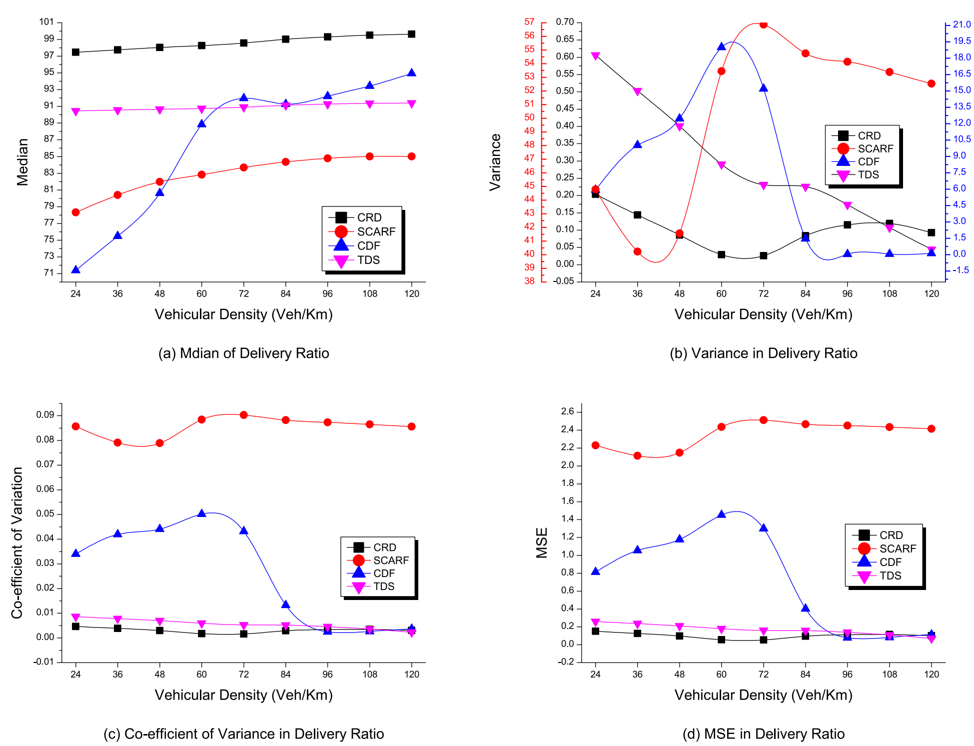
Fig. 14 indicates the performance in terms of transmission delay by varying vehicular speed and selfish nodes' density as mentioned above, with minimum and maximum values for the proposed scheme (7.4–7.8 ms) which is significantly less than SCARF's (8.8–19.1 ms), CDF's (25.1–49.6 ms), and TDS's (8.1–14.3 ms) by approximately 15.9%–59.16%, 70.52%–84.27%, and 8.64%–45.45% respectively. These results show that CRD disseminates the EWM much faster than its competitors regardless of the change in vehicular speed under varying selfish nodes' density. Similarly, to evaluate the degree of variation in transmission delay for each value vehicular speed under varying selfish nodes' density from 10% to 50%, Fig. 15 provides the minimum and maximum values of median (7.75–7.4) with a difference of 0.35, variance (0.028–0.001) with a difference of 0.027, co-efficient of variance (0.022–0.004) with a difference of 0.018, and MSE (0.056–0.011) with a difference of 0.045 for CRD as compared to SCARF's (12.1–11.4) with a difference of 0.7, (13.86–10.16) with a difference of 3.7, (0.28–0.26) with a difference of 0.2, and (1.24–1.06) with a difference of 0.18,



(a) CRD (b) SCARF

(c) CDF (d) TDS

Fig. 8. Delivery ratio for varying vehicular and selfish nodes' density.



(a) Mdian of Delivery Ratio (b) Variance in Delivery Ratio

(c) Co-efficient of Variance in Delivery Ratio (d) MSE in Delivery Ratio

Fig. 9. Comparative analysis in terms of delivery ratio and vehicular density.

CDF's (40.9–27) with a difference of 13.9, (26.48–0.7) with a difference of 25.77, (0.12–0.03) with a difference of 0.09, and (1.7–0.28) with a difference of 1.44, and TDS's (10.9–9.2) with a difference of 1.7, (3.46–0.44) with a difference of 3.02, (0.17–0.07) with a difference of 0.1, and (0.62–0.22) with a difference of 0.4, respectively. These results indicate that the performance of CRD remains very stable as compared to its competitors even if vehicular speed is varied significantly under a wide range of selfish nodes' density in the network.

6.2.3. Hop-count

Fig. 16 shows the number of hops required to deliver the EWM to a following node in the network by varying selfish nodes' density and vehicular densities. It can be observed that the proposed scheme

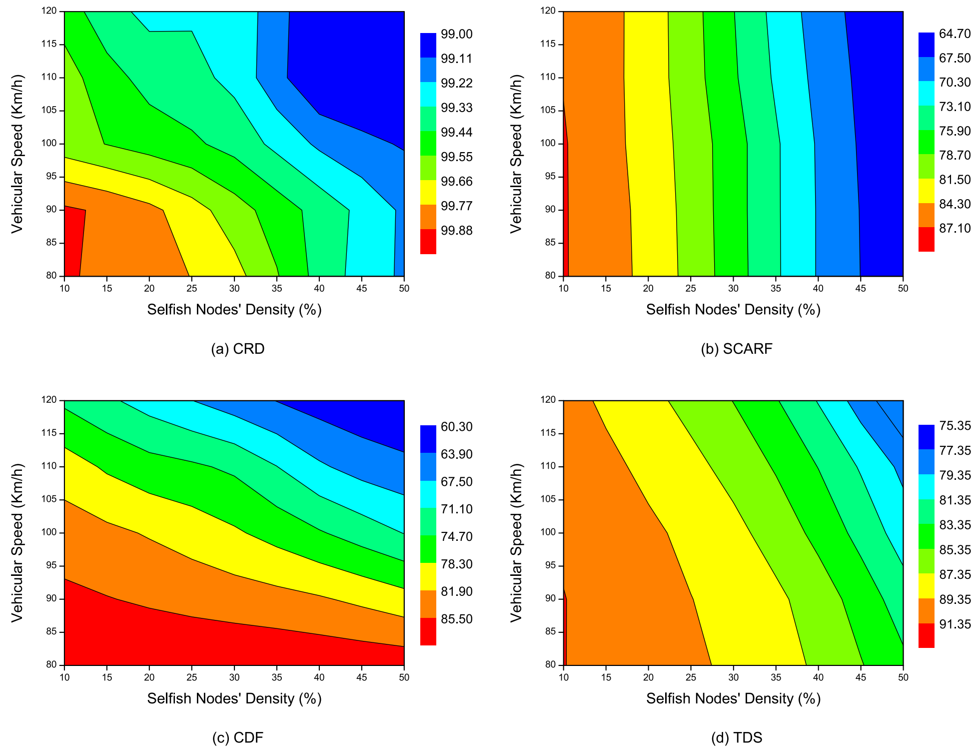


Fig. 10. Delivery ratio for varying vehicular speed and selfish nodes' density.

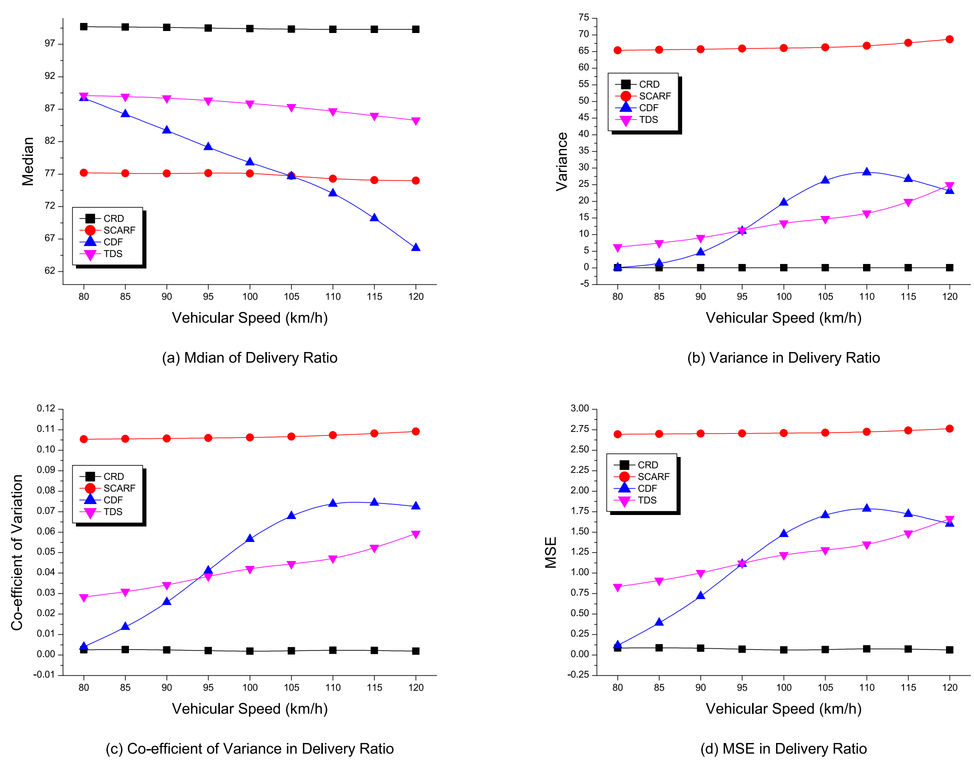


Fig. 11. Comparative analysis in terms of delivery ratio and vehicular speed.

achieves better results with average minimum and maximum values of hop-count (1–1.22) which is significantly less than SCARF's (1–3.6), CDF's (1.3–3), and TDS's (2–2.98) by approximately 0%–66.11%, 23.08%–59.33%, and 50%–59.06% respectively. These results show that CRD requires significantly fewer number of hops than its competitors regardless of the vehicular density under varying selfish nodes' density. Similarly, to evaluate the degree of variation in hop-count for each vehicular density value under varying selfish nodes' density, Fig. 17 shows the values of median (1–1.25) with a difference of 0.25, variance (0–0.007) with a difference of 0.007, co-efficient of variance (0–0.08) with a difference of 0.08, and MSE (0–0.028) with a difference of 0.028 for CRD as compared to SCARF's (1.4–2.98) with a difference

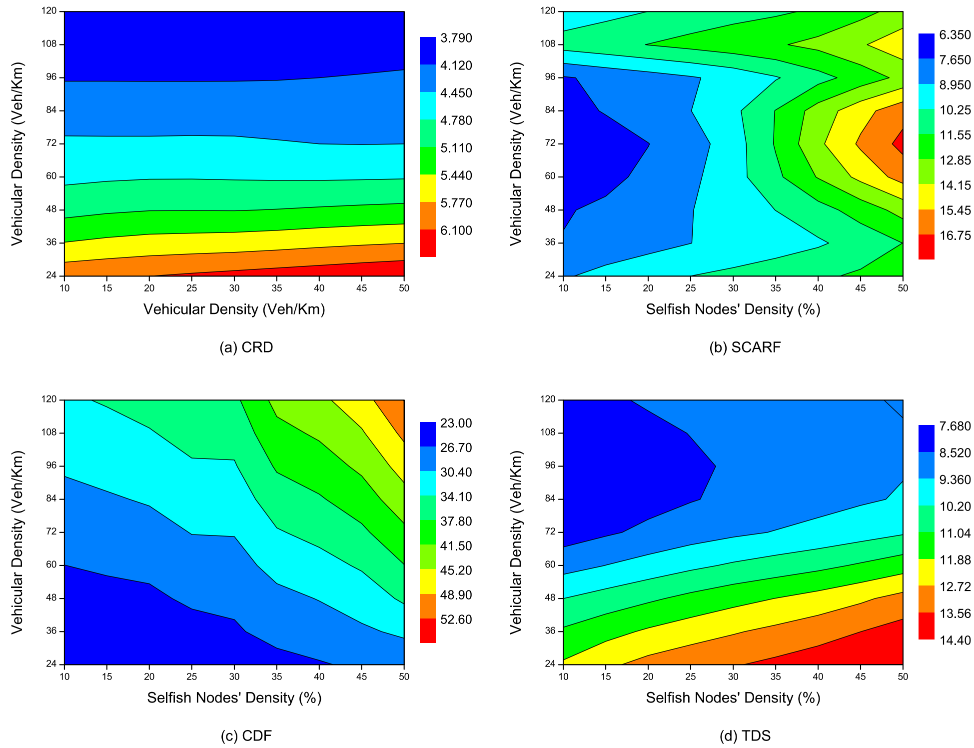


Fig. 12. Transmission delay for varying vehicular density and selfish nodes' density.

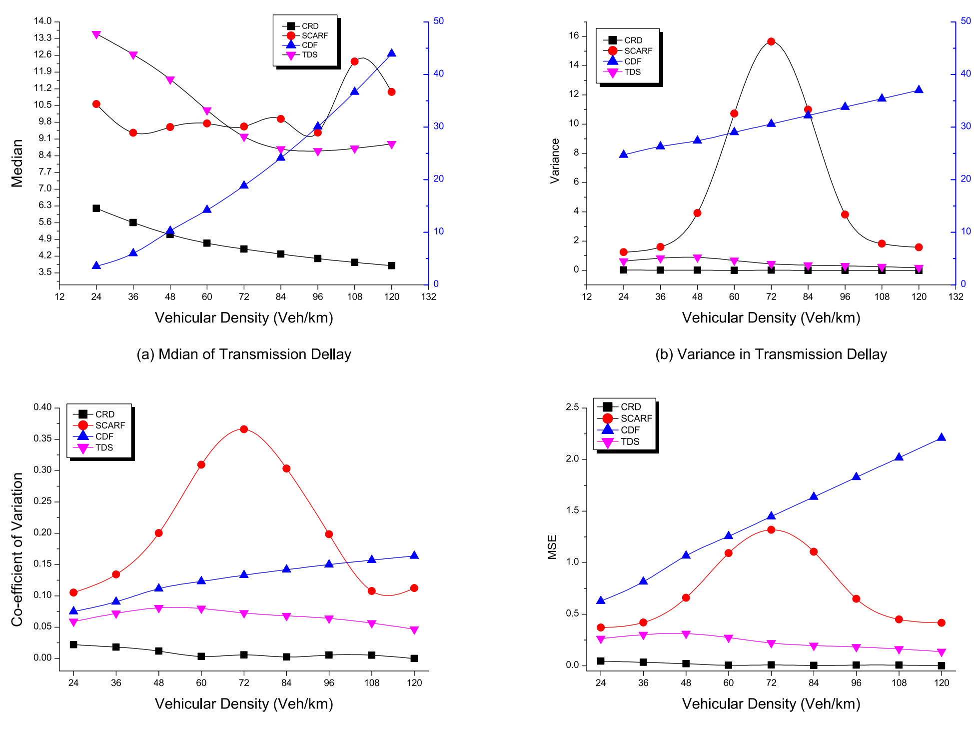


Fig. 13. Comparative analysis in terms of transmission delay and vehicular density.

of 1.58, (0.02–0.34) with a difference of 0.32, (0.06–0.35) with a difference of 0.29, and (0.04–0.19) with a difference of 0.15, CDF's (2–2.55) with a difference of 0.55, (0.009–0.283) with a difference of 0.27, (0.03–0.22) with a difference of 0.19, and (0.03–0.17) with a difference of 0.14, and TDS's (2.1–2.7) with a difference of 0.6, (0.001–0.127) with a difference of 0.13, (0.03–0.22) with a difference of 0.19, and (0.01–0.12) with a difference of 0.11 respectively. These results indicate that the proposed scheme is least affected by variation in vehicular and selfish nodes' density in terms of number of hops as compared to its competitors.

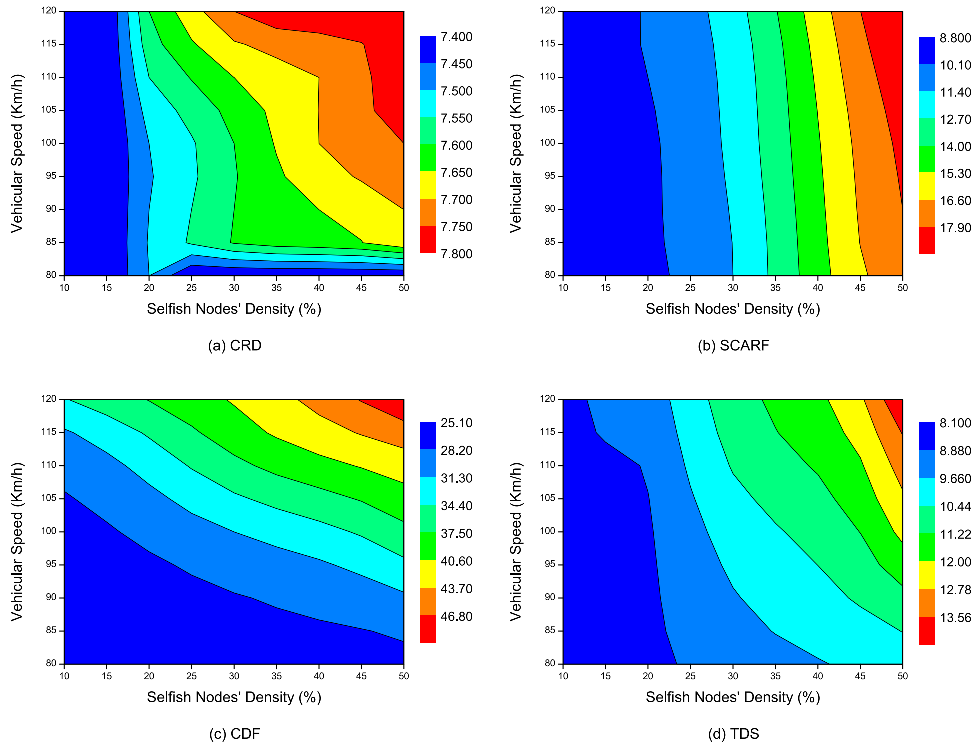


Fig. 14. Transmission delay for varying vehicular speed and selfish nodes' density.

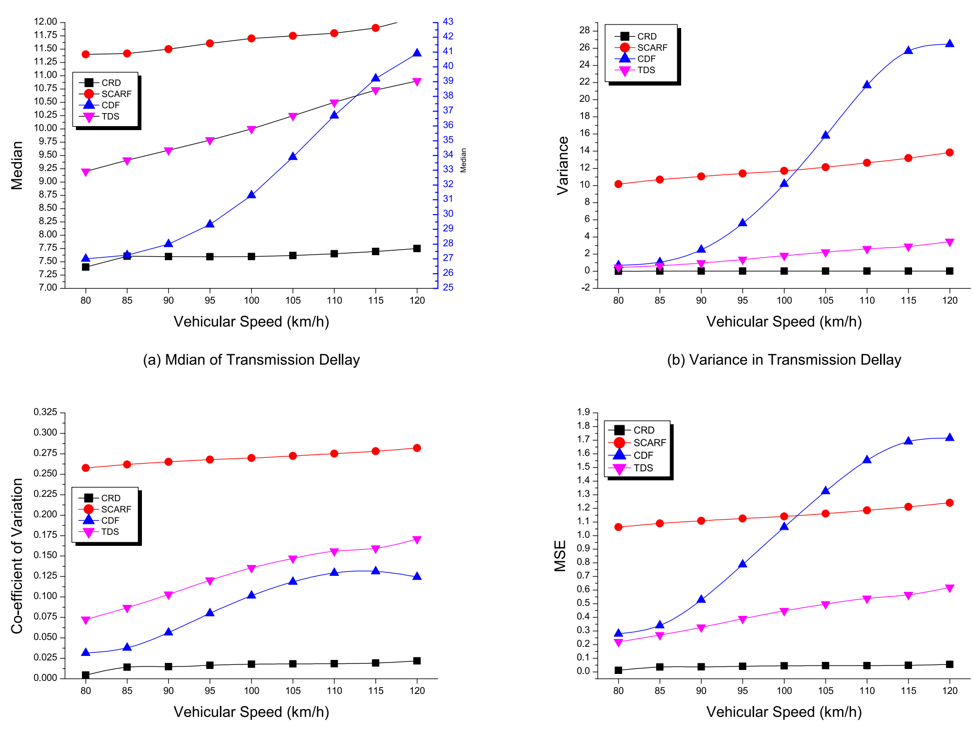
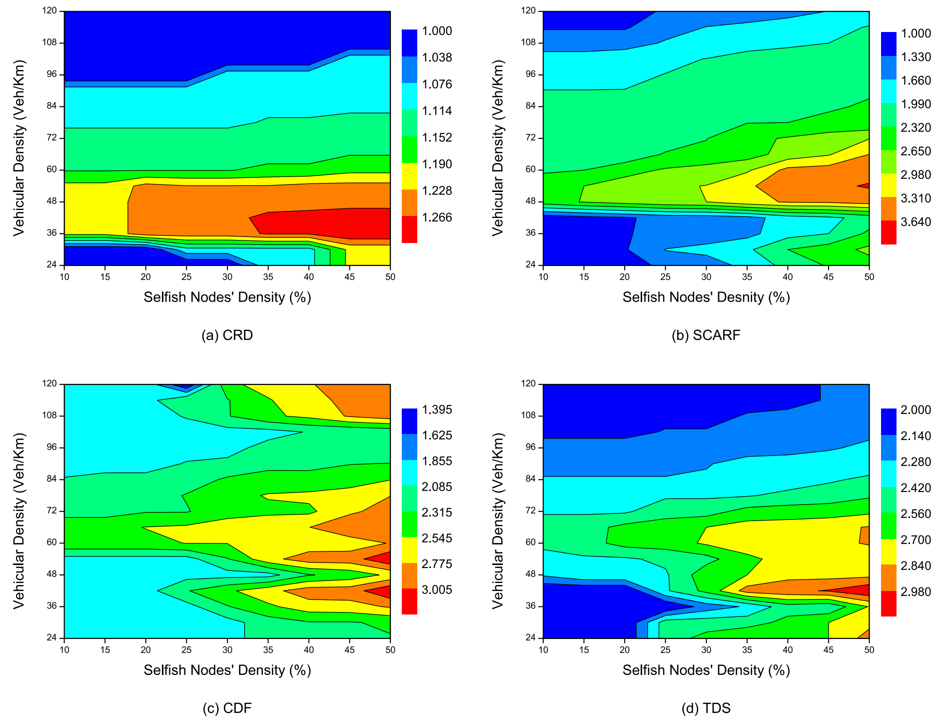


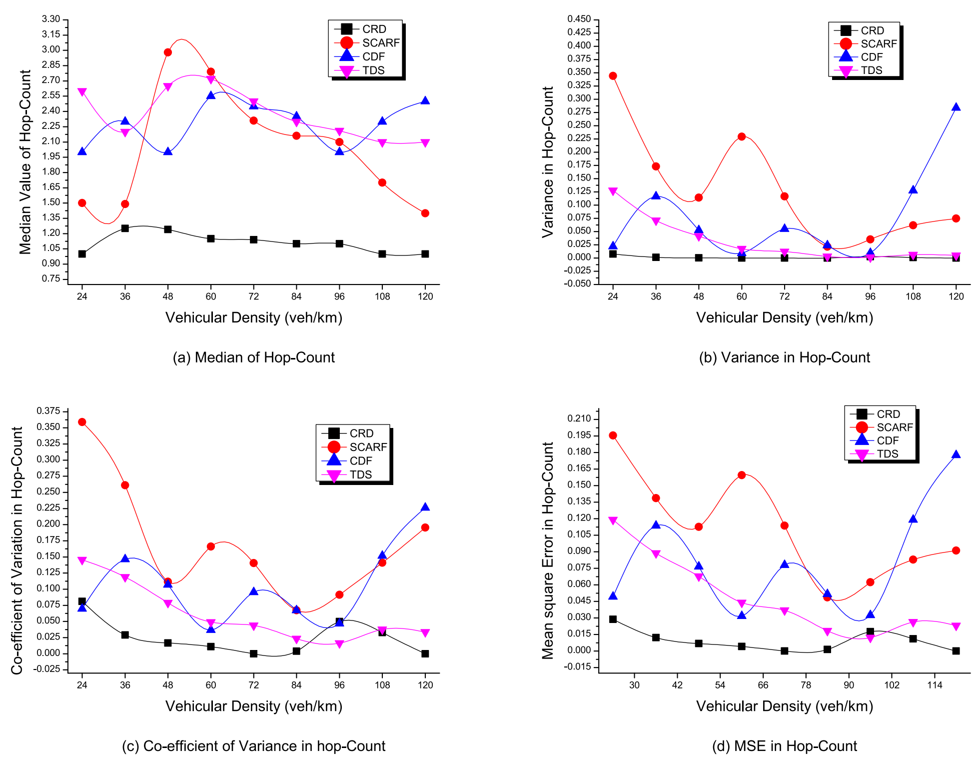
Fig. 15. Comparative analysis in terms of transmission delay and vehicular speed.

Fig. 18 illustrates the data of hop-count by varying vehicular speed and selfish nodes' density as mentioned above, with minimum and maximum values for the proposed scheme (1.14–1.2) which is significantly less than SCARF's (2.6–4.7), CDF's (3.5–6.8), and TDS's (3.1–4.2) by approximately 56.15%–74.47%, 67.43%–82.35%, and 62.23%–71.43% respectively. These results show that CRD requires significantly fewer number of hops than its competitors regardless of the change in vehicular speed under varying selfish nodes' density. Similarly, Fig. 19 evaluates the degree of variation in hop-count for each vehicular speed value under varying selfish nodes' density with minimum and



(a) CRD (b) SCARF (c) CDF (d) TDS

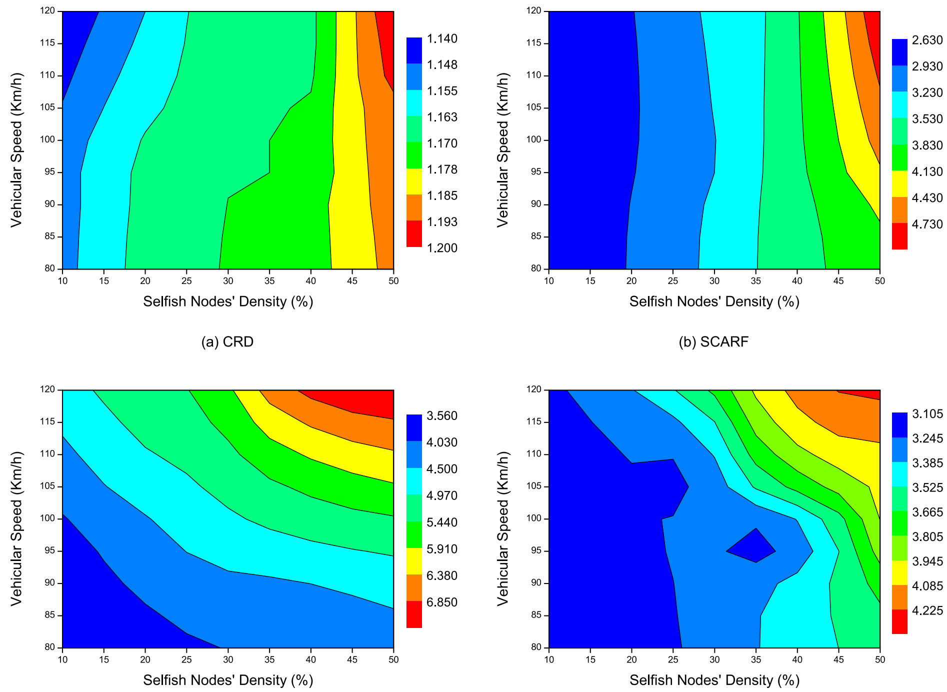
Fig. 16. Hop-count for varying vehicular density and selfish nodes' density.



(a) Median of Hop-Count (b) Variance in Hop-Count (c) Co-efficient of Variance in hop-Count (d) MSE in Hop-Count

Fig. 17. Comparative analysis in terms of hop-count and vehicular density.

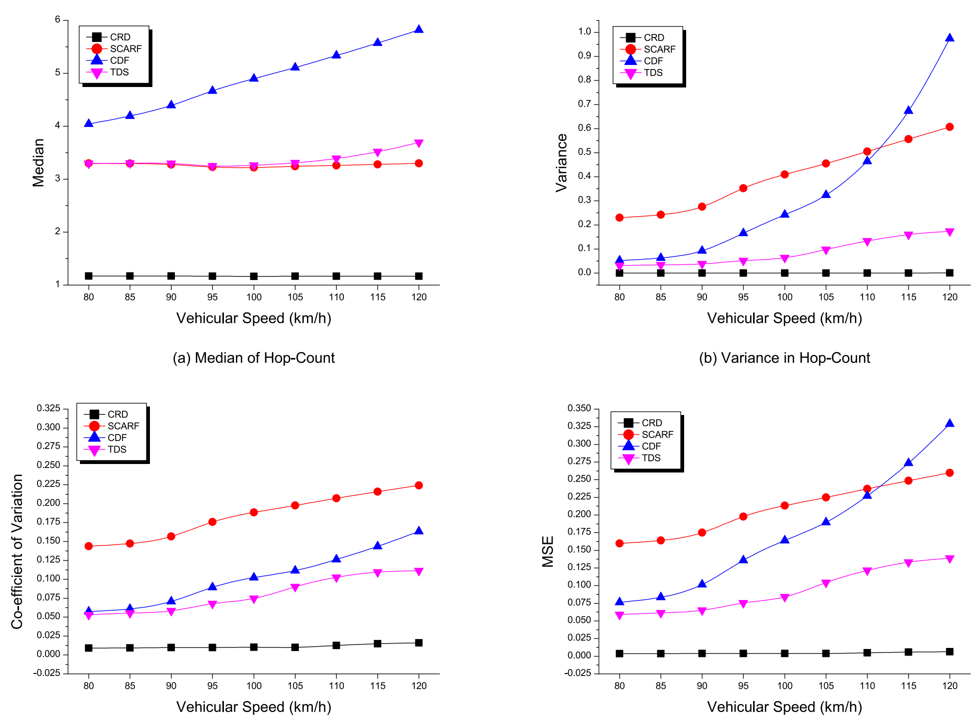
maximum values of median (1.16–1.17) with a difference of 0.01, variance (0.001–0.003) with a difference of 0.002, co-efficient of variance (0.009–0.015) with a difference of 0.006, and MSE (0.003–0.006) with a difference of 0.003 for CRD as compared to SCARF's (3.22–3.3) with a difference of 0.08, (0.23–0.6) with a difference of 0.37, (0.14–0.22) with a difference of 0.08, and (0.15–0.26) with a difference of 0.11, CDF's (4–5.9) with a difference of 1.9, (0.05–0.98) with a difference of 0.93, (0.06–0.16) with a difference of 0.1, and (0.07–0.33) with a difference of 0.26, and TDS's (3.25–3.69) with a difference of 0.44, (0.03–0.17) with a difference of 0.14, (0.05–0.11) with a difference of 0.06, and (0.06–0.13) with a difference of 0.07 respectively. These results indicate that the performance of CRD remains



(a) CRD (b) SCARF

(c) CDF (d) TDS

Fig. 18. Hop-count for varying vehicular speed and selfish nodes' density.



(a) Median of Hop-Count (b) Variance in Hop-Count

(c) Co-efficient of Variance in hop-Count (d) MSE in Hop-Count

Fig. 19. Comparative analysis in terms of hop-count and vehicular speed.

very stable as compared to its competitors even if vehicular speed is varied significantly under a wide range of selfish nodes' density in the network.

6.2.4. Total No. of transmissions

Fig. 20 shows the total number of transmitted messages in order to disseminate the EWM successfully throughout the network by varying vehicular (24 veh/km 120 veh/km) and selfish nodes' (10%–50%) densities. The results indicate that the proposed scheme has the least network overhead with average minimum and maximum values of number of transmissions (127–617) which is significantly less than SCARF's (224–805), CDF's (246–771), and TDS's (192–724) by approximately 43.3%–23.35%, 48.37%–19.97%, and 33.85%–14.78% respectively. These results show that CRD requires significantly fewer

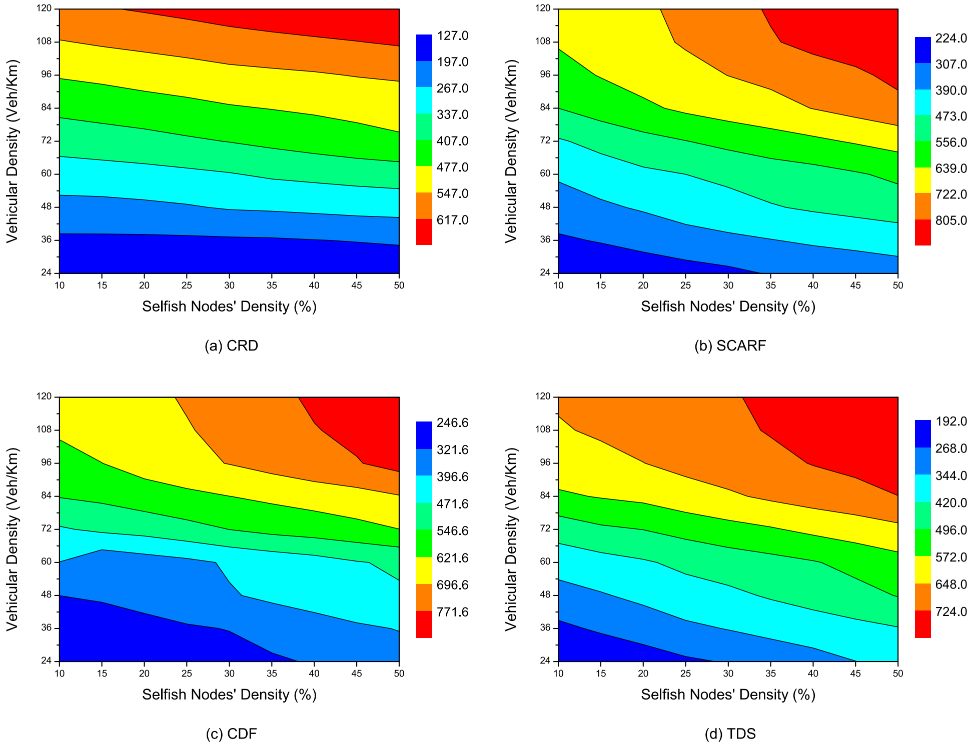


Fig. 20. No. of transmissions for varying vehicular density and selfish nodes' density.

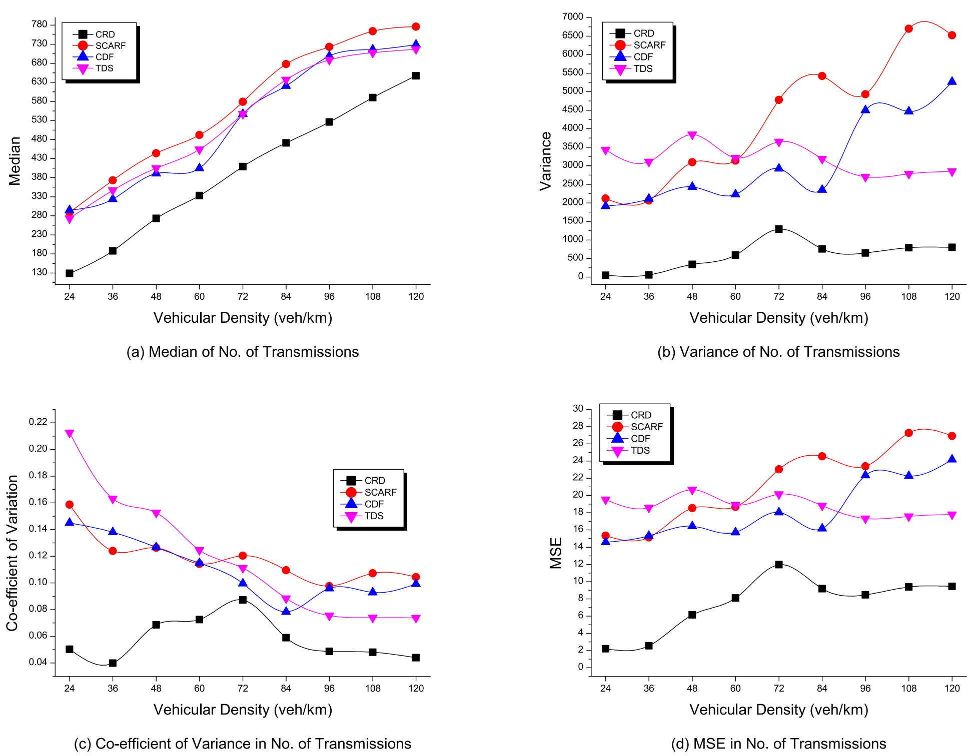


Fig. 21. Comparative analysis in terms of No. of transmissions and vehicular density.

total number of transmissions than its competitors regardless of the vehicular density under varying selfish nodes' density. Similarly, evaluating the degree of variation in network overhead for each vehicular density under varying selfish nodes' density, Fig. 21 shows the values of median (130–604) with a difference of 474, variance (0.7–2.45) with a difference of 1.75, co-efficient of variance (0.002–0.012) with a difference of 0.01, and MSE (0.28–0.5) with a difference of 0.22 for CRD as compared to SCARF's (260–716) with a difference of 456, (575–3948) with a difference of 3373, (0.06–0.09) with a difference of 0.03, and (8–21) with a difference of 13, CDF's (255–646) with a difference of 391, (45–76) with a difference of 31, (0.01–0.02) with a difference of 0.01, and (2.24–2.91) with a difference of 0.67, and TDS's (223–672)

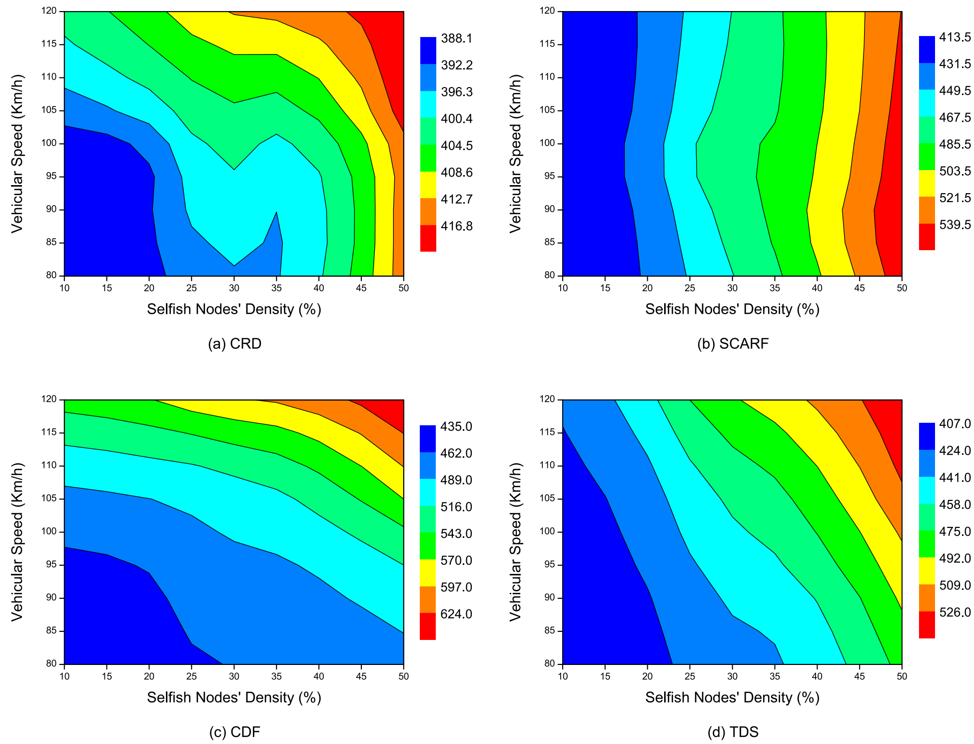


Fig. 22. No. of transmissions for varying vehicular density and selfish nodes' density.

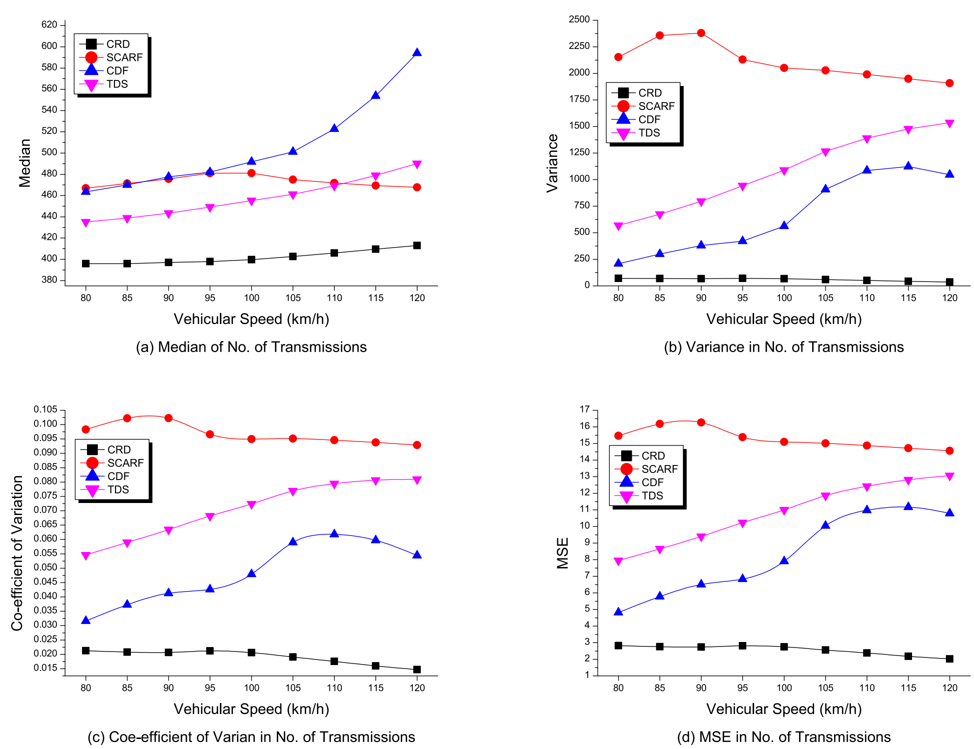


Fig. 23. Comparative analysis in terms of No. of transmissions and vehicular speed.

with a difference of 449, (313–489) with a difference of 176, (0.03–0.09) with a difference of 0.06, and (6–7.3) with a difference of 1.3 respectively. These results indicate that the proposed scheme is least affected by variation in vehicular and selfish nodes' density in terms of total number of transmissions as compared to its competitors.

Fig. 22 shows the total number of transmissions against varying vehicular speed and selfish nodes' density as mentioned above, with minimum and maximum values for the proposed scheme (388–416) which is significantly less than SCARF's (413–539), CDF's (435–624), and TDS's (407–526) by approximately 6%–22.82%, 10.8%–33.33%, 4.67%–20.91% respectively. These results show that CRD requires significantly fewer total number of transmissions than its competitors regardless of the change in vehicular speed under varying selfish nodes' density. Similarly, Fig. 23 evaluates the degree of variation in total number of transmissions for each vehicular speed value under varying selfish nodes' density with minimum and maximum values of median (395–413) with a difference of 18, variance (36.7–71.3) with a difference of 34.6, co-efficient of variance (0.014–0.102) with a difference of 0.089, and MSE (2–2.8) with a difference of 1.8 for CRD as compared to SCARF's (466–481) with a difference of 15, (1908–2380) with a difference of 472, (0.09–0.1025) with a difference of 0.0125, and (14.5–16.2) with a difference of 1.7, CDF's (463–594) with a difference of 131, (209–1122) with a difference of 913, (0.03–0.06) with a difference of 0.03, and (4.8–11.1) with a difference of 6.3, and TDS's (435–490) with a difference of 55, (568–1536) with a difference of 968, (0.05–0.08) with a difference of 0.03, and (7.9–13) with a difference of 5.1 respectively. These results indicate that the performance of CRD remains very stable as compared to its competitors even if vehicular speed is varied significantly under a wide range of selfish nodes' density in the network.

7. Conclusion

We highlighted the issues of connectivity, resource availability, and latency in conventional vehicular networks as well as limitations of spatio-temporal VSNs such as the presence of selfish nodes' and their adverse effects on dissemination of time-critical EWMs. For this purpose, we formulated VSN as a context-aware spatio-temporally evolving AMP and utilized the highly stable and ubiquitous connectivity oriented 5G network for communication instead of DSRC to achieve maximum accuracy in real-time scenarios. Then to mitigate the impact of selfishness, we proposed a social-intelligence based mechanism to differentiate socially-selfish and cooperative nodes by calculating a new tie-strength metric based on nodes' intimacy, similarity, social connectivity and social activity. Furthermore, a recursive evolutionary algorithm based mechanism is then presented such that the cooperative nodes gain higher reputation whereas selfish nodes are encouraged to improve their reputation to benefit from the system. In addition, we designed a real-time state-transition based dissemination algorithm for EWMs in a VSN environment by estimating the probability of most suitable nodes to become super-spreader. Finally, we performed extensive experimental and comparative analysis to assess the performance of the proposed scheme against state-of-the-art similar schemes in terms of delivery ratio, transmission delay, message overhead, and hop-count under varying network conditions such as vehicular density, speed, and density of selfish nodes. The obtained results have validated the efficacy of the proposed scheme in terms of scalability, robustness, and resilience to selfish nodes in the network.

In future, we plan to extend this work to include urban mobility models that require completely different road network topology. Moreover, reliable and real social information based dataset generation, especially in sparse networks, is still an open issue that can be explored to further analyze and enhance the performance of VSN based dissemination schemes. Similarly, privacy issues in AMPs still pose a great challenge to ensure trust of users in such solutions, which can be improved by exploring social network theory in conjunction with AI and cryptographic analysis. Further research in AI can reduce the algorithm complexity manifold. In addition, human involvement can be further reduced to improve the performance of such systems by adopting the AMPs based automatic solutions and 5G network based applications.

CRediT authorship contribution statement

Noor Ullah: Methodology, Software, Writing - original draft. **Xiangjie Kong:** Conceptualization, Supervision, Writing - review & editing. **Limei Lin:** Investigation, Writing - review & editing. **Mubarak Alrashoud:** Validation, Writing - review & editing. **Amr Tolba:** Writing - review & editing. **Feng Xia:** Supervision, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors extend their appreciation to the International Scientific Partnership Program ISPP at King Saud University for funding this research work through ISPP-78.

References

[1] O. Urmonov, H. Kim, A multi-hop data dissemination algorithm for vehicular communication, Computers 9 (2) (2020) 25.
[2] Y. Yang, Q. Liu, Z. Gao, X. Qiu, L. Rui, X. Li, A data dissemination mechanism for motorway environment in VANETs, EURASIP J. Wireless Commun. Networking 2015 (1) (2015) 93.
[3] S. Latif, S. Mahfooz, B. Jan, N. Ahmad, Y. Cao, M. Asif, A comparative study of scenario-driven multi-hop broadcast protocols for VANETs, Veh. Commun. 12 (2018) 88–109.
[4] A. Mchergui, T. Moulahi, B. Alaya, S. Nasri, A survey and comparative study of QoS aware broadcasting techniques in VANET, Telecommun. Syst. 66 (2) (2017) 253–281.
[5] L. Atzori, A. Floris, R. Girau, M. Nitti, G. Pau, Towards the implementation of the social internet of vehicles, Comput. Netw. 147 (2018) 132–145.
[6] A. Rahim, X. Kong, F. Xia, Z. Ning, N. Ullah, J. Wang, S.K. Das, Vehicular social networks: A survey, Pervasive Mob. Comput. 43 (2018) 96–113.
[7] B. Zhang, R. Tian, C. Li, Content dissemination and routing for vehicular social networks: A networking perspective, IEEE Wirel. Commun. (2019) 1–9.
[8] A.M. Vegni, V. Loscri, R. Petrolo, SCARF: A social-aware reliable forwarding technique for vehicular communications, in: Proceedings of the 3rd Workshop on Experiences with the Design and Implementation of Smart Objects, Association for Computing Machinery, New York, NY, USA, 2017, pp. 1–6.
[9] J. Li, Z. Ning, B. Jedari, F. Xia, I. Lee, A. Tolba, Geo-social distance-based data dissemination for socially aware networking, IEEE Access 4 (2016) 1444–1453.
[10] N. Ullah, X. Kong, L. Wan, H. Chen, Z. Wang, F. Xia, A social utility-based dissemination scheme for emergency warning messages in vehicular social networks, Comput. J. 61 (7) (2018) 971–986.
[11] N. Ullah, X. Kong, Z. Ning, A. Tolba, M. Alrashoud, F. Xia, Emergency warning messages dissemination in vehicular social networks: a trust based scheme, Veh. Commun. 22 (2020) 100199.
[12] G. Araniti, A. Orsino, L. Militano, L. Wang, A. Iera, Context-aware information diffusion for alerting messages in 5G mobile social networks, IEEE Internet Things J. 4 (2) (2017) 427–436.
[13] C. Campolo, A. Molinaro, A. Iera, A reference framework for social-enhanced vehicle-to-everything communications in 5G scenarios, Comput. Netw. 143 (2018) 140–152.
[14] T. Yang, X. Long, A.K. Sangaiah, Z. Zheng, C. Tong, Deep detection network for real-life traffic sign in vehicular networks, Comput. Netw. 136 (2018) 95–104.
[15] E.-K. Lee, M. Gerla, G. Pau, U. Lee, J.-H. Lim, Internet of vehicles: From intelligent grid to autonomous cars and vehicular fogs, Int. J. Distrib. Sens. Netw. 12 (9) (2016) 1550147716665500.
[16] X. Kong, F. Xia, J. Li, M. Hou, M. Li, Y. Xiang, A shared bus profiling scheme for smart cities based on heterogeneous mobile crowdsourced data, IEEE Trans. Ind. Inf. 16 (2) (2020) 1436–1444.
[17] C. Gao, H. Zhang, X. Chen, Y. Li, D. Jin, S. Chen, Impact of selfishness in device-to-device communication underlying cellular networks, IEEE Trans. Veh. Technol. 66 (10) (2017) 9338–9349.
[18] F. Xia, B. Jedari, L.T. Yang, J. Ma, R. Huang, A signaling game for uncertain data delivery in selfish mobile social networks, IEEE Trans. Comput. Soc. Syst. 3 (2) (2016) 100–112.
[19] B. Jedari, F. Xia, H. Chen, S.K. Das, A. Tolba, Z. AL-Makhadmeh, A social-based watchdog system to detect selfish nodes in opportunistic mobile networks, Future Gener. Comput. Syst. 92 (2019) 777–788.
[20] Z. Ning, L. Liu, F. Xia, B. Jedari, I. Lee, W. Zhang, CAIS: A copy adjustable incentive scheme in community-based socially aware networking, IEEE Trans. Veh. Technol. 66 (4) (2017) 3406–3419.
[21] A. Rahim, K. Ma, W. Zhao, A. Tolba, Z. Al-Makhadmeh, F. Xia, Cooperative data forwarding based on crowdsourcing in vehicular social networks, Pervasive Mob. Comput. 51 (2018) 43–55.
[22] A. Socievole, A. Caputo, F. De Rango, P. Fazio, Routing in mobile opportunistic social networks with selfish nodes, Wirel. Commun. Mobile Comput. 2019 (2019) 1–15.

[23] A. Bharath, P.P. Sivagurunathan, Collision Avoidance system in vehicular adhoc network utilizing dichotomized headway model, in: 2014 International Conference on Circuits, Power and Computing Technologies, ICCPCT-2014, 2014, pp. 1349–1353.
[24] F.D. Cunha, G.G. Maia, A.C. Viana, R.A. Mini, L.A. Villas, A.A. Loureiro, Socially inspired data dissemination for vehicular ad hoc networks, in: Proceedings of the 17th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, Association for Computing Machinery, New York, NY, USA, 2014, pp. 81–85.
[25] A. Dua, N. Kumar, S. Bawa, ReIDD: reliability-aware intelligent data dissemination protocol for broadcast storm problem in vehicular ad hoc networks, Telecommun. Syst. 64 (3) (2017) 439–458.
[26] A.M. Vegni, C. Souza, V. Loscri, E. Hernandez-Orallo, P. Manzoni, Data transmissions using hub nodes in vehicular social networks, IEEE Trans. Mob. Comput. (2019) 1.
[27] X. Lin, X. Li, Achieving efficient cooperative message authentication in vehicular ad hoc networks, IEEE Trans. Veh. Technol. 62 (7) (2013) 3339–3348.
[28] J. Guo, X. Li, Z. Liu, J. Ma, C. Yang, J. Zhang, D. Wu, TROVE: A context awareness trust model for VANETs using reinforcement learning, IEEE Internet Things J. (2020) 1.
[29] J. Miao, O. Hasan, S.B. Mokhtar, L. Brunie, K. Yim, An investigation on the unwillingness of nodes to participate in mobile delay tolerant network routing, Int. J. Inf. Manage. 33 (2) (2013) 252–262.
[30] B. Jedari, F. Xia, Z. Ning, A survey on human-centric communications in non-cooperative wireless relay networks, IEEE Commun. Surv. Tutor. 20 (2) (2018) 914–944.
[31] H. Gong, L. Yu, X. Zhang, Social contribution-based routing protocol for vehicular network with selfish nodes, Int. J. Distrib. Sens. Netw. 10 (4) (2014) 753024.
[32] A.M. Ahmed, X. Kong, L. Liu, F. Xia, S. Abolfazli, Z. Sanaei, A. Tolba, BoDMaS: Bio-inspired selfishness detection and mitigation in data management for ad-hoc social networks, Ad Hoc Netw. 55 (2017) 119–131, Self-organizing and Smart Protocols for Heterogeneous Ad hoc Networks.
[33] A. Rahim, T. Qiu, Z. Ning, J. Wang, N. Ullah, A. Tolba, F. Xia, Social acquaintance based routing in vehicular social networks, Future Gener. Comput. Syst. 93 (2019) 751–760.
[34] N. Basta, A. ElNahas, H.P. Grossmann, S. Abdennadher, A framework for social tie strength inference in vehicular social networks, in: 2019 Wireless Days (WD), 2019, pp. 1–8.
[35] X. Fan, L. Liu, M. Li, Z. Su, EigenTrustp++: Attack resilient trust management, in: 8th International Conference on Collaborative Computing: Networking, Applications and Worksharing, CollaborateCom, 2012, pp. 416–425.
[36] J. Lee, J.C. Oh, A model for recursive propagations of reputations in social networks, in: Proc. ASONAM ?3, Niagra Falls, ON, Canada, 2013, pp. 666–670.
[37] Z. Ning, F. Xia, N. Ullah, X. Kong, X. Hu, Vehicular social networks: Enabling smart mobility, IEEE Commun. Mag. 55 (5) (2017) 48–55.
[38] A.M. Vegni, V. Loscrí, P. Manzoni, Analysis of small-world features in vehicular social networks, in: 2019 16th IEEE Annual Consumer Communications Networking Conference, CCNC, 2019, pp. 1–2.
[39] A. Zheng, S. Luo, H. Xia, A meme propagation model to combine social affirmation with meme attractiveness and persistence, Internat. J. Modern Phys. C 27 (01) (2016) 1650002.
[40] X. Kong, F. Xia, Z. Ning, A. Rahim, Y. Cai, Z. Gao, J. Ma, Mobility dataset generation for vehicular social networks based on floating car data, IEEE Trans. Veh. Technol. 67 (5) (2018) 3874–3886.
[41] F. Xia, A. Rahim, X. Kong, M. Wang, Y. Cai, J. Wang, Modeling and analysis of large-scale urban mobility for green transportation, IEEE Trans. Ind. Inf. 14 (4) (2018) 1469–1481.

Noor Ullah received his Bachelor and Master degrees in Electrical Engineering with major in Communication from University of Engineering and Technology Peshawar, Pakistan. He is currently a Ph.D. candidate in Software Engineering, at School of Software, Dalian University of Technology, China. His research interests include but not limited to Vehicular Socials Networks, Mobile Social Computing and Intelligent Transportation System design.

Xiangjie Kong received the B.Sc. and Ph.D. degrees from Zhejiang University, Hangzhou, China. He is a Professor with College of Computer Science and Technology, Zhejiang University of Technology. Previously, he was an Associate Professor with the School of Software, Dalian University of Technology, China. He has published over 130 scientific papers in international journals and conferences (with over 100 indexed by ISI SCIE). His research interests include network science, mobile computing, and computational social science. He is a Senior Member of the IEEE and CCF and is a member of ACM.

Limei Lin received her Ph.D. and M.S. degrees in College of Mathematics and Informatics from Fujian Normal University in 2016 and 2013, respectively, and she was a visiting scholar in Montclair State University in 2015. She pursued her postdoctoral career in Fujian Agriculture and Forestry University from 2016 to 2018. She is currently an associate Professor in the College of Mathematics and Informatics, Fujian Normal University. Her research interests include interconnection networks, network security, fault diagnosis, fault tolerant computing, and bioinformatics. She has published over 40 papers in these areas. Specially, she has published over 10 papers in IEEE journals, including IEEE Transactions on Computers, IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Dependable and Secure Computing and IEEE Transactions on Reliability.

Mubarak Alrashoud received the Ph.D. degree in computer science from Ryerson University, Toronto, ON, Canada, in 2015, He is currently an Associate Professor and the Head of the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. His research interests include social network analysis, Vehicular Ad hoc Networks, Social Awareness networks and software engineering decision support.

Amr Tolba received the M.Sc. and Ph.D. degrees from the Department of Mathematics and Computer Science, Faculty of Science, Menoufia University, Egypt, in 2002 and 2006, respectively. He is currently an Associate Professor with the Faculty of Science, Menoufia University, Egypt. He is currently on leave from Menoufia University to the Department of Computer Science, Community College, King Saud University (KSU), Saudi Arabia. He has authored or coauthored over 60 scientific articles in top-ranked (ISI) international journals and conference proceedings. His main research interests include socially aware networks, vehicular ad hoc networks, the Internet of Things, intelligent systems, and cloud computing.

Feng Xia received the B.Sc. and Ph.D. degrees from Zhejiang University, Hangzhou, China. He is currently an Associate Professor and the Discipline Leader in School of Engineering, IT and Physical Sciences, Federation University Australia, Ballarat, VIC, Australia. Dr. Xia has published 2 books and over 300 scientific papers in international journals and conferences. His research interests include data science, social computing, and systems engineering. He is a Senior Member of the IEEE and ACM.